DOCUMENT RESUME

ED 447 473 CS 014 284

AUTHOR	Donahue, Patricia L.; Finnegan, Robert J.; Lutkus, Anthony D.; Allen, Nancy L.; Campbell, Jay R.
TITLE	The Nation's Report Card: Fourth-Grade Reading, 2000.
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-2001-499
PUB DATE	2001-04-00
NOTE	140p.; In collaboration with Dave Freund, Steve Isham, Laura Jerry, Youn-hee-Lim, Shari Santapau, Spence Swinton, Lois Worthington. For the highlights of this report, see CS 014 285; for a PowerPoint slide presentation related to this report, see CS 014 286.
AVAILABLE FROM	ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827 (Toll Free); Web site: http://www.ed.gov/pubs/edpubs.html. For full text: http://nces.ed.gov/nationsreportcard.
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Research (143)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Comparative Analysis; Grade 4; Intermediate Grades; *National Competency Tests; *Reading Achievement; Reading Research; Sex Differences; Standardized Tests; *Student Evaluation; Tables (Data); Test Results
IDENTIFIERS	*National Assessment of Educational Progress

ABSTRACT

This report presents the results of the 2000 NAEP (National Assessment of Educational Progress) fourth-grade reading assessment for the nation. Results in 2000 are compared to previous NAEP reading assessments. After an introduction, chapter 1 presents average scale scores and achievement level results for the nation. Chapter 2 presents average scale score and achievement level results for selected subgroups of the fourth-grade students. In Chapter 3, school and home contexts for learning are addressed. Chapter 4 discusses becoming a more inclusive national assessment. Major findings are: (1) the reading performance of the nation's fourth graders remained relatively stable across assessment years; (2) significant changes were evident at the upper and lower ends of the performance distribution--higher performing students made progress, and the score at the 10th percentile in 2000 was significantly lower than 1992; (3) in 2000, the percentage of fourth-grade students performing at or above the "basic" level was 63%, and performance at or above the "proficient" level was achieved by 32% of fourth graders; (4) female fourth graders had a higher average score than their male peers; (5) white and Asian/Pacific Islander students outperformed their black, Hispanic, and American Indian peers; (6) students in the Northeast and Central regions outperformed their counterparts in the Southeast and the West; (7) students in central city schools had a lower average score than their peers in urban fringe/large town and rural/small town locations; (8) students eligible for the free/reduced lunch program had a lower average score than students ineligible for that program; (9) students attending public schools had lower average scores than their peers attending nonpublic schools; (10) students who reported reading more pages daily in school and for homework had higher average scores than

Reproductions supplied by EDRS are the best that can be made
from the original document.

students reporting reading fewer pages daily; and (11) the average score for the nation was lower in the results that included the performance of students who needed and were provided with testing accommodations. Appendixes contain an overview of the procedures used; sample text and questions; and data. (RS)

Reproductions supplied by EDRS are the best that can be made
from the original document.

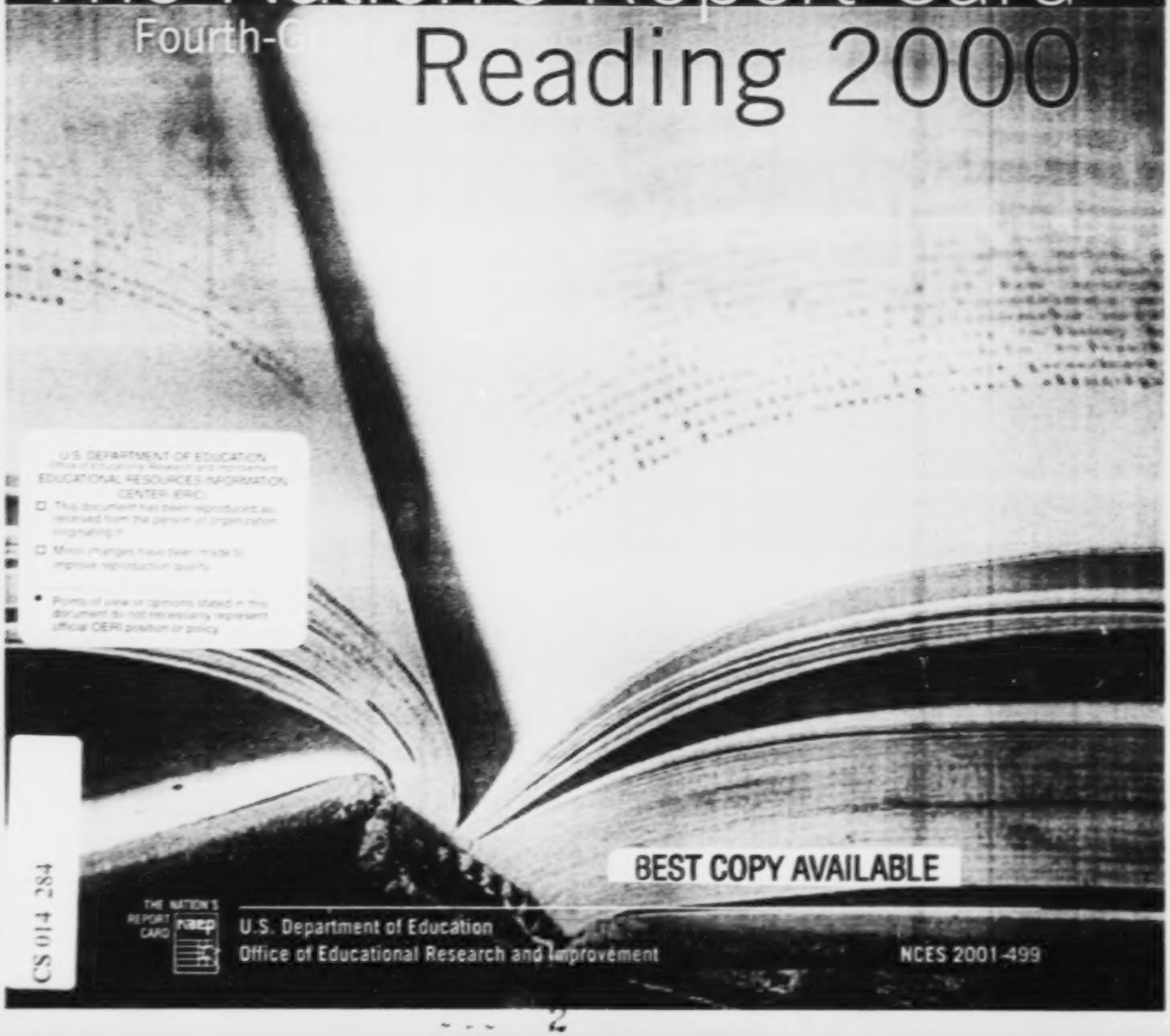
ED 447 473

National Center for Education Statistics

The Nation's Report Card

Fourth-G

Reading 2000



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS 014 284

THE NATION'S REPORT CARD naep

U.S. Department of Education
Office of Educational Research and Improvement

NCES 2001-499

What is The Nation's Report Card?

THE NATION'S REPORT CARD, the National Assessment of Educational Progress (NAEP), is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. Since 1969, assessments have been conducted periodically in reading, mathematics, science, writing, history, geography, and other fields. By making objective information on student performance available to policymakers at the national, state, and local levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education. Only information related to academic achievement is collected under this program. NAEP guarantees the privacy of individual students and their families.

NAEP is a congressionally mandated project of the National Center for Education Statistics, the U.S. Department of Education. The Commissioner of Education Statistics is responsible, by law, for carrying out the NAEP project through competitive awards to qualified organizations. NAEP reports directly to the Commissioner, who is also responsible for providing continuing reviews, including validation studies and solicitation of public comment, on NAEP's conduct and usefulness.

In 1988, Congress established the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP. The Board is responsible for selecting the subject areas to be assessed from among those included in the National Education Goals; for setting appropriate student performance levels; for developing assessment objectives and test specifications through a national consensus approach; for designing the assessment methodology; for developing guidelines for reporting and disseminating NAEP results; for developing standards and procedures for interstate, regional, and national comparisons; for determining the appropriateness of test items and ensuring they are free from bias; and for taking actions to improve the form and use of the National Assessment.

The National Assessment Governing Board

Mark D. Musick, Chair
President
Southern Regional Education Board
Atlanta, Georgia

Michael T. Nettles, Vice Chair
Professor of Education
University of Michigan
Ann Arbor, Michigan

Moses Barnes
Secondary School Principal
Fort Lauderdale, Florida

Melanie A. Campbell
Fourth-Grade Teacher
Topeka, Kansas

Honorable Wilmer S. Cody
Former Commissioner of Education
State of Kentucky
Frankfort, Kentucky

Daniel A. Domenech
Superintendent of Schools
Fairfax County Public Schools
Fairfax, Virginia

Edward Donley
Former Chairman
Air Products & Chemicals, Inc.
Allentown, Pennsylvania

Thomas H. Fisher
Director
Student Assessment Services
Florida Department of Education
Tallahassee, Florida

Edward H. Haertel
Professor, School of Education
Stanford University
Stanford, California

Juanita Haugen
Local School Board Member
Pleasanton, California

Honorable Nancy Kopp
State Legislator
Annapolis, Maryland

Honorable Ronnie Musgrove
Governor of Mississippi
Jackson, Mississippi

Roy M. Nageak, Sr.
First Vice-Chair
Alaska Board of Education and
Early Development
Barrow, Alaska

Debra Paulson
Eighth-Grade Mathematics Teacher
El Paso, Texas

Honorable Jo Ann Pottorff
State Legislator
Wichita, Kansas

Diane Ravitch
Research Professor
New York University
New York, New York

Sister Lourdes Sheehan, R.S.M.
Secretary for Education
United States Catholic Conference
Washington, DC

John H. Stevens
Executive Director
Texas Business and Education
Coalition
Austin, Texas

Adam Urbanski
President
Rochester Teachers Association
Rochester, New York

Migdania D. Vega
Principal
Coral Way Elementary Bilingual
School
Miami, Florida

Deborah Voltz
Assistant Professor
Department of Special Education
University of Louisville
Louisville, Kentucky

Honorable Michael E. Ward
State Superintendent of Public
Instruction
North Carolina Public Schools
Raleigh, North Carolina

Marilyn A. Whirry
Twelfth-Grade English Teacher
Manhattan Beach, California

Dennie Palmer Wolf
Senior Research Associate
Harvard University
Graduate School of Education
Cambridge, Massachusetts

(Ex-Officio)
Assistant Secretary of Education
Office of Educational Research and
Improvement
U.S. Department of Education
Washington, DC

Roy Truby
Executive Director, NAGB
Washington, DC

National Center for
Education Statistics

The Nation's Report Card

Fourth-Grade Reading 2000

Patricia L. Donahue

Robert J. Finnegan

Anthony D. Lutkus

Nancy L. Allen

Jay R. Campbell

In collaboration with

Dave Freund	**Shari Santapau**
Steve Isham	**Spence Swinton**
Laura Jerry	**Lois Worthington**
Youn-hee Lim	

April 2001

U.S. Department of Education
Office of Educational Research and Improvement

NCES 2001–499

U.S. Department of Education
Rod Paige
Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

April 2001

SUGGESTED CITATION
U.S. Department of Education. Office of Educational Research and Improvement. National Center for Education Statistics. ***The** Nation's Report Card: Fourth-Grade Reading 2000*, NCES 2001–499, by P. L. Donahue, R. J. Finnegan, A. D. Lutkus, N.L. Allen, and J. R. Campbell. Washington, DC: 2001.

FOR MORE INFORMATION
Content contact:
Sheida White
202–502–7473

To obtain single copies of this report, limited number of copies available, or ordering information on other U.S. Department of Education products, call toll free 1–877–4ED PUBS (877–433–7827), or write:

Education Publications Center (ED Pubs)
U.S. Department of Education
P.O. Box 1398
Jessup, MD 20794–1398

TTY/TDD 1–877–576–7734
FAX 301–470–1244

Online ordering via the Internet: *http://www.ed.gov/pubs/edpubs.html*
Copies also are available in alternate formats upon request.
This report also is available on the World Wide Web: *http://nces.ed.gov/nationsreportcard*

The work upon which this publication is based was performed for the National Center for Education Statistics, Office of Educational Research and Improvement, by Educational Testing Service.

Table of Contents

The
Nation's
Report
Card

Executive Summary ix

Introduction 1

Overview of the 2000 National Assessment of Educational Progress (NAEP) 2
Framework for the 1992, 1994, 1998, and 2000 Assessments 2
The Reading Assessment Instruments 4
Description of School and Student Samples 5
Reporting the Assessment Results 5
The Setting of Achievement Levels 6
The Developmental Status of Achievement Levels 7
Interpreting NAEP Results 10
This Report 10

Chapter 1
Average Scale Score and Achievement Level Results for the Nation 11

Average Scale Score Results 12
Achievement Level Results 13
Sample Assessment Questions and Student Responses 16
Map of Selected Item Descriptions 21
Summary 23

Chapter 2
Average Scale Score and Achievement Level Results for Selected Subgroups of 4th-Grade Students 25

Gender 26
Race/Ethnicity 28
Trends in Scale Score Differences Between Selected Subgroups 33
Region of the Country 34
Type of Location 37
Eligibility for the Free/Reduced-Price Lunch Program 39
Type of School 41
Summary 45

Chapter 3

School and Home Contexts for Learning 47
- Reading in and for School 48
- Literacy with Family and Friends 54
- Summary 62

Chapter 4

Becoming a More Inclusive National Assessment 63
- Two Sets of 2000 NAEP Reading Results 64
- Results for the Nation: Accommodations Not Permitted and Accommodations Permitted 68
- Results by Gender: Accommodations Not Permitted and Accommodations Permitted 70
- Results by Race/Ethnicity: Accommodations Not Permitted and Accommodations Permitted 70
- Summary 72

Appendix A

Overview of Procedures Used for the NAEP 2000 Grade 4 Reading Assessment 73
- The NAEP 2000 Grade 4 Reading Assessment 73
- The Assessment Design 75
- National Sample 76
- Students with Disabilities (SD) and Limited English Proficient (LEP) Students 79
- Data Collection and Scoring 83
- Data Analysis and IRT Scaling 83
- Item Mapping Procedures 86
- Weighting and Variance Estimation 87
- Drawing Inferences from the Results 88
- Analyzing Group Differences in Averages and Percentages 89
- Conducting Multiple Tests 91
- NAEP Reporting Groups 93
- Cautions in Interpretations 96

Appendix B

Sample Text and Questions from the NAEP 2000 Reading Assessment 97

Appendix C

Data Appendix 103

Acknowledgments 125

Figures and Tables

Figure i.1
Descriptions of the three reading purposes specified in the NAEP framework 3

Figure i.2
Descriptions of the four reading stances specified in the NAEP framework 4

Figure i.3
Policy definitions of the three achievement levels 7

Figure 1.1
Average fourth-grade reading scale scores: 1992–2000 12

Figure 1.2
Fourth-grade reading scale score percentiles: 1992–2000 13

Figure 1.3
Descriptions of reading achievement for the levels at grade 4 14

Figure 1.4
Percentage of fourth-grade students at or above reading achievement levels
and within each achievement level range: 1992–2000 15

Table 1.1
Sample question 1 results (multiple-choice) 17

Table 1.2
Sample question 2 results (short constructed-response) 18

Table 1.3
Sample question 3 results (extended constructed-response) 19

Figure 1.5
Map of selected item descriptions on the
National Assessment of Educational Progress
reading scale for grade 4 22

Figure 1.6
Average scale score, percentile scores,
and achievement level results: 1992–2000 23

Figure 2.1
Average fourth-grade reading scale scores
for male and female students: 1992–2000 26

Figure 2.2
Percentages of fourth-grade male and female students at or above
reading achievement levels and within each achievement level range:
1992–2000 27

Figure 2.3
Average fourth-grade reading scale scores by race/ethnicity: 1992–2000 29

Figure 2.4
Percentages of fourth-grade students at or above reading achievement levels
and within each achievement level range by race/ethnicity: 1992–2000 30

Figure 2.5
Differences in average fourth-grade reading scale scores
by gender and race/ethnicity: 1992–2000 33

Figure 2.6
Average fourth-grade reading scale scores
by region of the country: 1992–2000 34

Figure 2.7
Percentages of fourth-grade students at or above reading achievement levels
and within each achievement level range by region of the country: 1992–2000 35

Table 2.1
Average fourth-grade reading scale scores
by school's type of location: 2000 37

Figure 2.8
Percentages of fourth-grade students at or above reading achievement levels
and within each achievement level range by school's type of location: 2000 38

Figure 2.9
Average fourth-grade reading scale scores by student eligibility
for the free/reduced-price lunch program: 1998–2000 39

Figure 2.10
Percentages of fourth-grade students at or above reading achievement levels
and within each achievement level range by student eligibility for the
free/reduced-price lunch program: 1998–2000 40

Figure 2.11
Average fourth-grade reading scale scores by type of school: 1992–2000 42

Figure 2.12
Percentages of fourth-grade students at or above reading achievement levels
and within each achievement level range by type of school: 1992–2000 43

Figure 2.13
Summary of scale scores and achievement levels: 1992-2000 46

Table 3.1
Students' reports on the number of pages read each day
in school and for homework: 1992–2000 48

Table 3.2
Student's reports on the amount of time spent doing
homework each day: 1992–2000 49

Table 3.3
Students' reports on how often they write long answers to
questions on tests or assignments that involved reading: 1992–2000 51

Table 3.4
Students' reports on how often their teachers help them
break words into parts and help them understand new words: 1998–2000 53

Table 3.5
Students' reports on how often they read for fun
on their own time: 1992–2000 55

Table 3.6
Students' reports on how often they discuss their studies at home
and talk about reading with their family and friends: 1992–2000 57

Table 3.7
Students' reports on the number of different types of reading materials in the home: 1992–2000 59

Table 3.8
Students' reports on the amount of time spent watching television each day: 1992–2000 61

Figure 4.1
The two sets of NAEP results based on a split-sample design 67

Table 4.1
Average score by type of results: 1998 and 2000 68

Table 4.2
Percentages of students attaining reading achievement levels by type of results: 1998 and 2000 69

Figure A.1
1992, 1994, 1998, and 2000 NAEP framework—aspects of grade 4 reading literacy 74

Table A.1
Target and actual percentage distribution of questions by reading purpose, 2000 NAEP grade 4 reading assessment 75

Table A.2
Target and actual percentage distribution of questions by reading stance, 2000 NAEP grade 4 reading assessment 75

Table A.3
Grade 4 national student sample sizes 77

Table A.4
NAEP 2000 school and student participation rates for the nation: Grade 4 public schools, nonpublic schools, and combined 78

Table A.5
Students with disabilities and limited English proficient students in the NAEP reading assessment: Grade 4 national samples where accommodations were not permitted: 1992–2000 80

Table A.6
Students with disabilities and limited English proficient students in the NAEP reading assessment: Grade 4 national samples where accommodations were permitted for identified students: 1998 and 2000 81

Table A.7
SD and LEP students assessed with accommodations, in the NAEP reading assessment: Grade 4 national samples where accommodations were permitted for identified students, public and nonpublic schools combined by type of accommodation: 1998 and 2000 82

Table A.8
FDR comparisons of average reading scale scores for race/ethnicity groups: 1994 and 2000 91

Figure A.2
States included in the four NAEP regions 93

Executive Summary

The Nation's Report Card

Reading Scale Score and Achievement Level Results for the Nation

Results for Student Subgroups

School and Home Contexts for Learning

Transitioning to a More Inclusive NAEP

The National Assessment of Educational Progress (NAEP) is the nation's only federally mandated survey of student achievement in various subject areas. Authorized by Congress and administered by the National Center for Education Statistics in the U.S. Department of Education, NAEP regularly reports to the public on the educational progress of students in grades 4, 8, and 12. In 2000, NAEP conducted a national reading assessment of fourth-grade students.

This report presents the results of the 2000 NAEP fourth-grade reading assessment for the nation. Results in 2000 are compared to results of previous NAEP reading assessments. Students' performance on the assessment is described in terms of average scores on a 0–500 scale, and in terms of the percentage of students attaining three achievement levels: *Basic*, *Proficient*, and *Advanced*. The achievement levels are performance standards adopted by the National Assessment Governing Board (NAGB) as part of its statutory responsibilities. They are collective judgments of what students should know and be able to do.

As provided by law, the Commissioner of Education Statistics, upon review of a congressionally mandated evaluation of NAEP, determined that the achievement levels are to be considered developmental and should be interpreted and used with caution. However, both the Acting Commissioner and the Board believe these performance standards are useful for understanding trends in student achievement. They have been widely used by national and state officials, including the National Education Goals Panel, as a common yardstick of academic performance.

In addition to providing average scores and achievement level performance in reading for the nation's fourth-graders, this report provides results for subgroups of fourth-grade students defined by various background and contextual characteristics. A summary of major findings from the 2000 NAEP reading assessment is presented on the following pages.

Reading Scale Score and Achievement Level Results for the Nation

The reading performance of the nation's fourth-graders has remained relatively stable across assessment years. In 2000, the national average scale score of 217 was similar to that in 1992.

Although the national average scale score has remained relatively stable, significant changes are evident at the upper and lower ends of the performance distribution. Higher performing students have made progress: scores at the 75th and 90th percentiles in 2000 were significantly higher than 1992. In contrast, the score at the 10th percentile in 2000 was significantly lower than 1992.

In 2000, the percentage of fourth-grade students performing at or above the *Basic* level of reading achievement was 63 percent. Performance at or above the *Proficient* level—the level identified by NAGB as the level that all students should reach—was achieved by 32 percent of fourth-graders. The highest level of performance, the *Advanced* level, was achieved by 8 percent of fourth-graders.

In 2000, the percentages of fourth-graders performing at or above *Proficient* and at *Advanced* were higher than in 1992.

Results for Student Subgroups

Gender

- In 2000, female fourth-grade students had a higher average score than their male peers. The scale-score gap between males and females widened since 1998.
- The percentage of females at or above the *Proficient* level exceeded that of males.
- The percentage of female fourth-graders at or above the *Proficient* level in 2000 was higher than in 1992.

Race/Ethnicity

- In 2000, white and Asian/Pacific Islander students outperformed their black, Hispanic, and American Indian peers.
- A significant increase was observed in the average scale score of Asian/Pacific Islander students, whose 2000 score was higher than in 1992. The 2000 average score of black students was significantly higher in comparison to 1994.
- The percentages of white and Asian/Pacific Islander students at or above the *Proficient* level exceeded that of other racial/ethnic groups.
- Only among Asian/Pacific Islander students was an increase observed in the percentage at or above *Proficient* since 1992.

Region

- The 2000 results by region show fourth-grade students in the Northeast and Central regions outperforming their counterparts in the Southeast and the West.
- Among students in the Northeast, the average scale score in 2000 was higher in comparison to 1994.

- Students in the Northeast and Central regions had higher percentages of students at or above the *Proficient* level than the Southeast. The Northeast region had a higher percentage of students at or above *Proficient* than the West.

Type of Location

- Fourth-grade students in central city schools had a lower average score in 2000 than their peers who attended schools in urban fringe/large town and rural/small town locations.
- Comparisons of achievement level results between locations show a lower percentage of central city students at or above the *Proficient* level than their peers in other types of location.

Eligibility for Free/Reduced-Price Lunch

- In 2000, students who were eligible for the free/reduced-price lunch program had a lower average score than students who were ineligible for the program.
- Achievement level results also show lower performance among students eligible for the program. In 2000, 14 percent of eligible students performed at or above the *Proficient* level in comparison to 41 percent of noneligible students.

Type of School

- Consistent with past NAEP reading assessments, the 2000 results indicated that students attending public schools had lower average reading scale scores than their peers attending nonpublic schools.
- A lower percentage of public school students performed at or above the *Proficient* level in comparison to nonpublic school students.

School and Home Contexts for Learning

Pages Read in School and for Homework

- Fourth-graders who reported reading more pages daily in school and for homework had higher average scores than students who reported reading fewer pages daily.
- The 2000 results indicate that more fourth-grade students are reading eleven or more pages in school and for homework on a daily basis than in 1992 and 1994.

Time Spent Doing Homework

- Fourth-graders who reported spending a moderate amount of time on homework—one-half hour or one hour daily—had higher average scores than students who reported that they spent more than an hour or that they either did not have or did not do homework.
- The percentage of students in 2000 who reported that they do not have homework was lower in comparison to 1992 and 1994.

Writing about Reading

- Fourth-graders who reported writing long answers to questions on tests and assignments that involved reading on a weekly or monthly basis had higher average scores than students who reported doing so once or twice a year, or never or hardly ever.
- The 2000 assessment reports by fourth-graders indicate an increase in the frequency of writing about reading on a weekly basis in comparison to 1994.

Teachers' Help with Words

- Fourth-grade students who reported that their teachers never or hardly ever helped them break words into parts scored higher than their peers who reported receiving such help daily or weekly.
- Fourth-graders who reported that their teachers helped them understand new words on a weekly or monthly basis scored higher than those who reported receiving this help daily or never or hardly ever.

Reading for Fun

- Students who reported reading for fun on their own time every day had higher average scores than students who reported reading for fun less frequently.
- In 2000, 75 percent of fourth-grade students reported reading for fun on their own time at least weekly.

Discussing Studies and Talking about Reading

- Students who reported discussing their studies at home daily, weekly, or monthly had higher average scores than students who reported never or hardly ever having such discussions.
- Students who reported talking about their reading with family and friends on a weekly basis had a higher average score than students who reported engaging in such conversations daily, monthly, or never or hardly ever.
- In 2000, 61 percent of fourth-grade students reported talking about their reading with family or friends at least weekly.

Reading Materials in the Home

- The average score for students who reported having all four types of reading materials (books, magazines, newspapers, encyclopedia) in their home was higher than those who reported having fewer reading materials.
- In 2000, a lower percentage of students reported having all four types of reading materials in the home in comparison to 1994.

Time Spent Watching Television

- Students who reported watching three or fewer hours of television each day outperformed students who reported watching more television.
- In 2000, the percentages of students who reported watching four or more hours of television daily has decreased since 1994 and the percentages of students reporting watching three hours or less has increased since 1994.

Transitioning to a More Inclusive NAEP

- A second set of results from the 2000 NAEP reading assessment represents the performance of students when testing accommodations are permitted for special-needs students.
- A comparison of the two sets of results show that the average score for the nation was lower in the results that included the performance of students that needed and were provided with testing accommodations.
- A comparison of the two sets of results for Hispanic students show that their average score was lower in the results that included the performance of students that needed and were provided with testing accommodations.

Introduction

At the start of a new century, the importance of providing all students with an education that ensures them the means for personal, social, and economic growth remains as central to the American enterprise as it has always been. It may be, however, that the challenge inherent in this enterprise has never been greater. While the country had long anticipated technological advances, during the last decade progress seemed to accelerate and daily reality was transformed in many ways. One such transformation involved the way people read.

Traditional images of reading mainly represented an individual with a book in hand, but contemporary images more frequently include the image of an individual reading from a computer screen. Many of these individuals are our nation's students. Growing up in a time of unprecedented access to information, these students need the skills to read widely and critically in order to successfully navigate the variety of information available to them. Whether the text is print- or electronic-based, the act of reading in an increasingly complex world requires acute understanding and interpretation of all that is read.

The centrality of reading has long been recognized in the curriculum of our schools for its importance in shaping both personal selves and participants in a democratic society. The need to ensure the literate participation of our society's future citizens underlies the need to monitor our students' reading achievement.

Focus Questions

What is the NAEP reading assessment?

How does the NAEP reading assessment measure and report student progress?

Introduction Contents

Overview

Reading Framework

Reading Assessment

School and Student Samples

Reporting Results

NAEP Achievement Levels

Interpreting NAEP Results

This Report

Overview of the 2000 National Assessment of Educational Progress (NAEP)

In 1969, the National Assessment of Educational Progress was authorized by Congress to collect, analyze, and report reliable and valuable information about what American students know and can do in core subject areas. Since that time, in what has come to be referred to as the long-term trend assessment, NAEP has assessed public and nonpublic school students who are 9, 13, and 17 years old. Since 1990, the more recently developed assessments referred to as the main NAEP, have also assessed public and nonpublic school students in grades 4, 8, and 12. In 2000, student performance in mathematics and science was assessed at all three grades, and student performance in reading was assessed at grade 4 only.

All NAEP assessments are based on content frameworks developed through a national consensus process. The 2000 NAEP reading assessment was the fourth administration of an assessment based on *The NAEP Reading Framework*.[1] In 1992, 1994, and 1998, the NAEP reading assessment was administered to national samples of fourth-, eighth-, and twelfth-graders. In 1992 and 1994, the reading assessment was also administered to samples of fourth-graders participating in the state-by-state assessment, and in 1998 the state-by-state assessment included samples of fourth- and eighth-graders. Although the National Assessment Governing Board schedule calls for a reading assessment every four years, with the next full-scale reading assessment planned for 2002, an assessment at grade four only was authorized for 2000. The administration of the reading assessment in 2000 was in response to heightened concern that all children learn to read by the end of third grade. In order to closely monitor early reading achievement, the assessment was given to a national sample of fourth-graders. No state-by-state reading results were collected in 2000.

This report describes the results of the 2000 NAEP reading assessment at grade 4, and compares results in 2000 to fourth-grade performance in 1992, 1994, and 1998. The comparisons will focus on 2000 results in relation to earlier results. Comparisons of 1998 to 1994 and of 1994 to 1992 were made in previous report cards and therefore are neither highlighted nor discussed in this report.[2] Comparisons across assessment years are possible because the assessments were developed under the same framework and share a common set of reading tasks and because the populations of students were sampled and assessed using comparable procedures.

Framework for the 1992, 1994, 1998, and 2000 Assessments

The NAEP Reading Framework[3] provided the operational specifications and theoretical basis for developing all the NAEP reading assessments administered since 1992. The result of a national consensus effort, the *NAEP Reading Framework* reflects the ideas of many individuals involved and interested in reading education. This consensus effort was managed by the Council of Chief State School Officers (CCSSO) under the direction of the National Assessment Governing Board (NAGB).

---

[1] National Assessment Governing Board. *Reading framework for the National Assesment of Educational Progress: 1992–2000*. Washington, DC: Author.

[2] Donahue, P.L., Voekl, K.E., Campbell, J.R., & Mazzeo, J. (1999). *NAEP 1998 reading report card for the nation and the states*. Washington, DC: National Center for Education Statistics.

Campbell, J.R., Donahue, P.L., Reese, C.M., & Phillips, G.W. (1996). *NAEP 1994 reading report card for the nation and the states*. Washington, DC: National Center for Education Statistics.

[3] National Assessment Governing Board. *Reading framework for the National Assessment of Educational Progress 1992–2000*. Washington, DC: Author. [Also available online at http:www.nagb.org/pubs/92-2000read/toc.html]

The framework reflects research that views the act of reading as a dynamic and interactive process involving the reader, the text, and the context of the reading experience. For example, readers may employ different strategies depending on the text's structure and their purpose for reading it.[4]

Recognizing that readers vary their approach according to their purpose for reading, the framework specifies three purposes for reading to be assessed: reading for literary experience, reading to gain information, and reading to perform a task. All three purposes are assessed at grades 8 and 12, but reading to perform a task is not assessed at grade 4. The three purposes for reading as specified in the framework are described in figure i.1.

Figure i.1 **Descriptions of the three reading purposes specified in the NAEP framework**

Reading Purposes	
Reading for literary experience	Reading for literary experience entails the reading of various literary texts to enlarge our experience of human events and emotions, and to enhance both our appreciation of the world and how it is depicted through language. Literary texts used in the NAEP reading assessment included adventure stories, science fiction, and folktales.
Reading to gain information	When reading to gain information, readers are usually focused on a specific topic or point of reference. They are trying to understand and retain the text information. Informative texts used in the NAEP reading assessment included science articles, primary and secondary historical sources, sections of textbook chapters, essays, and speeches.
Reading to perform a task	Reading to perform a task involves reading various types of materials for the purpose of applying the information or directions to complete a specific task. As such, readers must focus on how they will actually use the information. The materials used to assess this purpose in the NAEP reading assessment included classified advertisements, directions for completing various projects, and a tax form.[Not assessed at grade 4.]

SOURCE: National Assessment Governing Board. *Reading framework for the National Assessment of Educational Progress: 1992–2000.*

---

[4] Anderson, R.C. & Pearson, P.D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P.D. Pearson (Ed.), *Handbook of reading research* (pp. 255–292). New York: Longman.

Pressley, M. & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ruddell, R.B. & Unrau, N.J. (1994). Reading as a meaning-construction process: The reader, the text, and the teacher. In R. B. Ruddell, M.R. Ruddell, & H. Singer (Eds.) *Theoretical models and processes of reading* (pp. 864–894). Newark, DE: International Reading Association.

Taylor, B.M. (1992). Text structure, comprehension, and recall. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 220–235). Newark, DE: International Reading Association.

The framework also specifies four types of reading processes—referred to as "reading stances"—that characterize ways readers respond to text. The four stances are: Initial Understanding, Developing an Interpretation, Personal Reflection and Response, and Critical Stance. The reading stances represent the changing ways readers position themselves in relation to a text, with each way contributing to the comprehension of it.[3] The stances are not intended to be indicative of hierarchical reading skills, but rather to represent aspects of reading processes that occur at any reading ability level. The four reading stances as specified in the framework are described in figure i.2.

The Reading Assessment Instruments

As the only federally mandated ongoing assessment of student reading achievement on a national scale, it is imperative that the NAEP assessment reflects the framework and expert perspectives and opinions about reading comprehension and its measurement. To that end, the assessment development process involves stages and processes of review by teachers and teacher educators as well as by state officials and measurement experts. All components of the assessment are evaluated for curricular relevance, developmental appropriateness, and fairness concerns.

Figure i.2 **Descriptions of the four reading stances specified in the NAEP framework**

Reading Stances

Initial understanding:	**preliminary consideration of the text as a whole** Readers are asked to consider the whole text in demonstrating an overall understanding of its meaning and function.
Developing an interpretation:	**discerning connections and relationships among ideas within the text** Readers are asked to build upon their initial impressions to develop a more thorough understanding of the text and the interrelationship of its parts.
Personal reflection and response:	**relating personal knowledge to text ideas** Readers are asked to describe how ideas in the text confirm, contradict, or compare with prior knowledge and experiences.
Critical stance:	**standing apart from the text to consider it objectively** Readers are asked to consider how the author conveys information, expresses ideas or feelings, and communicates a message.

SOURCE: National Assessment Governing Board. *Reading framework for the National Assessment of Educational Progress: 1992–2000.*

---

[3] Langer, J.A. (1990). The process of understanding: Reading for literary and informative purposes. *Research in the teaching of English, 24*(3), 229–259.

The reading passages used in the NAEP reading assessment are drawn from the types of books and magazines that students might encounter in or out of school. In striving for authentic reading experiences, the reading materials are neither abridged nor written for the assessment. They are reprinted in test booklets in a format that replicates as closely as possible their original publication.

At grade 4, all test booklets contain two 25-minute sections, each containing a reading passage and a set of approximately 10 comprehension questions. The questions are presented in both multiple-choice and constructed-response formats. At least half of the questions are constructed-response, which allow students to write out their own answers and explain and support their ideas. Constructed-response questions were of two types: short, requiring a one or two sentence answer, and extended, requiring a paragraph or full-page response.

The 2000 reading assessment at fourth grade used 8 different reading passages and comprised a total of 81 questions: 35 multiple-choice, 38 short constructed-response (scored according to a two- or three-level scoring rubric), and 8 extended constructed-response (scored according to a four-level rubric). The greater proportion of student response time was spent answering the constructed-response questions.

Description of School and Student Samples

The NAEP 2000 reading assessment was administered to fourth-graders at the national level. There was no state-by-state reading assessment in 2000. The findings in this report pertain to all fourth-graders in the nation, both those who do not need accommodations as well as those who are provided accommodations for their disability or limited English proficiency. The national results presented in the first three chapters of this report are based on a nationally representative sample of fourth-graders who can be meaningfully assessed without accommodations. Chapter 4 presents national results for a representative sample that includes the performance of students who needed and were provided with accommodations.

Each selected school that participated in the assessment and each student assessed represent a portion of the population of interest. For information on sample sizes and participation rates, see appendix A.

Reporting the Assessment Results

The results of student performance on the NAEP reading assessment are presented in two ways: as average scores on the NAEP composite reading scale, and in terms of the percentage of students attaining NAEP reading achievement levels. The average scale scores represent students' performance on the assessment. The achievement levels represent how that performance measured up against set expectations for achievement. The average scale scores represent what students *know and can do.* The achievement level results indicate the degree to which student performance meets expectations of what they *should know and be able to do.*

Average scale score results are presented on the NAEP reading composite scale, which ranges from 0-500. Students' responses on the NAEP 2000 reading assessment were analyzed to determine the percentages of students responding correctly to each multiple-choice question and the percentages of students responding at each score level for the constructed-response questions. The analysis entails summarizing the results on separate subscales for each reading purpose, then combining the separate scales to form a single composite reading scale. The relative contribution of each reading purpose at grade 4, as specified in the reading framework, is 55 percent for reading for literary experience and 45 percent for reading to gain information. (A full description of NAEP scales and scaling procedures can be found in the forthcoming *NAEP 2000 Technical Report.*)

Achievement level results are presented in terms of reading achievement levels as authorized by the NAEP legislation and adopted by the National Assessment Governing Board. For each grade tested, NAGB has adopted three achievement levels: *Basic, Proficient,* and *Advanced.* For reporting purposes, the achievement level cut scores are placed on the reading scale, resulting in four ranges: below *Basic, Basic, Proficient,* and *Advanced.*

The Setting of Achievement Levels

The 1988 NAEP legislation that created the National Assessment Governing Board directed the Board to identify "appropriate achievement goals...for each subject area" that NAEP measures.[6] The 1994 NAEP reauthorization reaffirmed many of the Board's statutory responsibilities, including "developing appropriate student performance standards for each age and grade in each subject area to be tested under the National Assessment."[7] In order to follow this directive and achieve the mandate of the 1988 statute "to improve the form and use of NAEP results," the Board undertook the development of student performance standards (called "achievement levels"). Since 1990, the Board has adopted achievement levels in mathematics, reading, U.S. history, world geography, science, writing, and civics.

The Board defined three levels for each grade: *Basic, Proficient,* and *Advanced.* The *Basic* level denotes partial mastery of the knowledge and skills that are fundamental for proficient work at a given grade. The *Proficient* level represents solid academic performance. Students reaching this level demonstrate competency over challenging subject matter. The *Advanced* level signifies superior performance at a given grade. For each grade, the levels are cumulative; that is, abilities achieved at the *Proficient* level presume mastery of abilities associated with the *Basic* level, and attainment of the *Advanced* level presumes mastery of both the *Basic* and *Proficient* levels. Figure i.3 presents the policy definitions of the achievement levels that apply across grades and subject areas. (Specific descriptions of reading achievement for the levels at grade 4 are presented in chapter 1.) Adopting three levels of achievement for each grade signals the importance of looking at more than one standard of performance. The Board believes, however, that all students should reach the *Proficient* level; the *Basic* level is not the desired goal, but rather represents partial mastery that is a step toward *Proficient.*

The achievement levels in this report were adopted by the Board based on a

---

[6] Public Law 100–297. (1988). National Assessment of Educational Progress Improvement Act (20 USC 1221). Washington, DC.

[7] Public Law 103–382. (1994). Improving America's Schools Act (20 USC 9010). Washington, DC.

Figure i.3 **Policy definitions of the three achievement levels**

Achievement Levels

Basic	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.
Proficient	This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.
Advanced	This level signifies superior performance.

SOURCE: National Assessment Governing Board.

standard-setting process designed and conducted under a contract with ACT. To develop these levels, ACT convened a cross section of educators and interested citizens from across the nation and asked them to judge what students should know and be able to do relative to a body of content reflected in the NAEP assessment framework for reading. This achievement level setting process was reviewed by an array of individuals including policymakers, representatives of professional organizations, teachers, parents, and other members of the general public. Prior to adopting these levels of student achievement, NAGB engaged a large number of persons to comment on the recommended levels and to review the results.

The results of the achievement level setting process, after NAGB approval, became a set of achievement level descriptions and a set of achievement level cut points on the 0-500 NAEP reading scale. The cut points are the scores that define the boundaries between below *Basic*, *Basic*, *Proficient*, and *Advanced* performance at grades 4, 8, and 12. The Board established these reading achievement levels in 1992 based upon the reading content framework; these levels were used for the 1992, 1994, 1998, and 2000 reading assessments.

The Developmental Status of Achievement Levels

The 1994 NAEP reauthorization law requires that the achievement levels be used on a developmental basis until the Commissioner of Education Statistics determines that the achievement levels are "reasonable, valid, and informative to the public."[8] Until that determination is made, the law requires the Commissioner and the Board to state clearly the developmental status of the achievement levels in all NAEP reports.

In 1993, the first of several congressionally mandated evaluations of the achievement level setting process concluded that the procedures used to set the achievement levels were flawed and that the percentage of students at or above any particular achievement level cut point may be underestimated.[9] Others have critiqued these evaluations, asserting that the weight of the empirical evidence does not support such conclusions.[10]

---

[8] The Improving America's Schools Act of 1994 (20 USC 9010) requires that the Commissioner base his determination on a congressionally mandated evaluation by one or more nationally recognized evaluation organizations, such as the National Academy of Education or the National Academy of Science.

In response to the evaluations and critiques, NAGB conducted an additional study of the 1992 reading achievement levels before deciding to use the 1992 reading achievement levels for reporting 1994 NAEP results.[11] When reviewing the findings of this study, the National Academy of Education (NAE) Panel expressed concern about what it saw as a "confirmatory bias" in the study and about the inability of this study to "address the panel's perception that the levels had been set too high." [12] In 1997, the NAE Panel summarized its concerns with interpreting NAEP results based on the achievement levels as follows:

> *First, the potential instability of the levels may interfere with the accurate portrayal of trends. Second, the perception that few American students are attaining the higher standards we have set for them may deflect attention to the wrong aspects of education reform. The public has indicated its interest in benchmarking against international standards, yet it is noteworthy that when American students performed very well on a 1991 international reading assessment, these results were discounted because they were contradicted by poor performance against the possibly flawed NAEP reading achievement levels in the following year.*[13]

The NAE Panel report recommended "that the current achievement levels be abandoned by the end of the century and replaced by new standards...." The National Center for Education Statistics and the National Assessment Governing Board have sought and continue to seek new and better ways to set performance standards on NAEP. [14] For example, NCES and NAGB jointly sponsored a national conference on standard setting in large-scale assessments, which explored many issues related to standard setting.[15] Although new directions were presented and discussed, a proven alternative to the current process has not yet been identified. The Commissioner of Education Statistics and the Board continue to call on the research community to assist in finding ways to improve standard setting for reporting NAEP results.

---

[9] United States General Accounting Office. (1993). *Education achievement standards: NAGB's approach yields misleading interpretations*. U.S. General Accounting Office Report to Congressional Requestors. Washington, DC: Author.

National Academy of Education. (1993). *Setting performance standards for achievement: A report of the National Academy of Education Panel on the evaluations of the NAEP Trial State Assessment: An evaluation of the 1992 achievement levels.* Stanford, CA: Author.

[10] Cizek, G. (1993). *Reactions to National Academy of Education report.* Washington, DC: National Assessment Governing Board.

Kane, M. (1993). *Comments on the NAE evaluation of the NAGB achievement levels.* Washington, DC: National Assessment Governing Board.

[11] American College Testing. (1995). *NAEP reading revisited: An evaluation of the 1992 achievement level descriptions.* Washington, DC: National Assessment Governing Board.

[12] National Academy of Education. (1996). Reading achievement levels. In *Quality and utility: The 1994 Trial State Assessment in reading. The fourth report of the National Academy of Education Panel on the evaluation of the NAEP Trial State Assessment.* Stanford, CA: Author.

[13] National Academy of Education. (1997). *Assessment in transition: Monitoring the nation's educational progress* (p. 99). Mountain View, CA: Author.

[14] Reckase, Mark. D. (2000). *The evolution of the NAEP achievement levels setting process: A summary of the research and development efforts conducted by ACT.* Iowa City, IA: ACT, Inc.

[15] National Assessment Governing Board and National Center for Education Statistics. (1995). *Proceedings of the joint conference on standard setting for large-scale assessments of the National Assessment Governing Board (NAGB) and the National Center for Education Statistics (NCES).* Washington, DC: Government Printing Office.

The most recent congressionally mandated evaluation conducted by the National Academy of Sciences (NAS) relied on prior studies of achievement levels, rather than carrying out new evaluations, on the grounds that the process has not changed substantially since the initial problems were identified. Instead, the NAS Panel studied the development of the 1996 science achievement levels. The NAS Panel basically concurred with earlier congressionally mandated studies. The Panel concluded that "NAEP's current achievement level setting procedures remain fundamentally flawed. The judgment tasks are difficult and confusing; raters' judgments of different item types are internally inconsistent; appropriate validity evidence for the cut scores is lacking; and the process has produced unreasonable results."[16]

The NAS Panel accepted the continuing use of achievement levels in reporting NAEP results on a developmental basis, until such time as better procedures can be developed. Specifically, the NAS Panel concluded that "…tracking changes in the percentages of students performing at or above those cut scores (or in fact, any selected cut scores) can be of use in describing changes in student performance over time."[17]

The National Assessment Governing Board urges all who are concerned about student performance levels to recognize that the use of these achievement levels is a developing process and is subject to various interpretations. The Board and the Acting Commissioner believe that the achievement levels are useful for reporting trends in the educational achievement of students in the United States.[18] In fact, achievement level results have been used in reports by the President of the United States, the Secretary of Education, state governors, legislators, and members of Congress. The National Education Goals Panel and government leaders in the nation and in more than 40 states use these results in their annual reports.

However, based on the congressionally mandated evaluations so far, the Commissioner agrees with the National Academy's recommendation that caution needs to be exercised in the use of the current achievement levels. Therefore, the Commissioner concludes that these achievement levels should continue to be considered developmental and should continue to be interpreted and used with caution.

---

[16] Pellegrino, J.W., Jones, L.R., & Mitchell, K.J. (Eds.). (1998). *Grading the nation's report card: evaluating NAEP and transforming the assessment of educational progress.* Committee on the Evaluation of National Assessments of Educational Progress, Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education, National Research Council. (p.182). Washington, DC: National Academy Press.

[17] *Ibid.*, page 176.

[18] Forsyth, Robert A. (2000). A description of the standard-setting procedures used by three standardized test publishers. In *Student performance standards on the National Assessment of Educational Progress: Affirmations and improvements.* Washington, DC: National Assessment Governing Board.

Nellhaus, Jeffrey M. (2000). States with NAEP-like performance standards. In *Student performance standards on the National Assessment of Educational Progress: Affirmations and improvements.* Washington, DC: National Assessment Governing Board.

Interpreting NAEP Results

The average scores and percentages presented in this report are estimates because they are based on representative samples of students rather than on each individual student in the population(s). As such, the results are subject to a measure of uncertainty, reflected in the standard error of the estimates. The standard errors for the estimated scale scores and percentages in this report are provided in appendix C.

The differences between scale scores and between percentages discussed in the following chapters take into account the standard errors associated with the estimates. Comparisons are based on statistical tests that consider both the magnitude of the difference between the group average scores or percentages and the standard errors of those statistics. Throughout this report, differences between scores or between percentages are pointed out only when they are significant from a statistical perspective. All differences reported are significant at the .05 level with appropriate adjustments for multiple comparisons. The term significant is not intended to imply a judgment about the absolute magnitude of the educational relevance of the differences. It is intended to identify statistically dependable population differences to help inform dialogue among policymakers, educators, and the public.

Readers are cautioned against interpreting NAEP results in a causal sense. Inferences related to subgroup performance or to the effectiveness of public and nonpublic schools, for example, should take into consideration the many socioeconomic and educational factors that may also impact reading performance.

This Report

This report describes fourth-grade students' performance on the NAEP 2000 reading assessment and compares the performance to previous NAEP assessment results. Chapter 1 presents overall scale score and achievement level results and also provides a more delineated view by showing scores at percentiles across the performance distribution. Chapter 2 examines assessment results for subgroups of students by gender, race/ethnicity, region of the country, type of location, type of school, and by eligibility for the free/reduced-price lunch program. In chapter 3, students' home and school experiences as portrayed by their responses to NAEP background questions provide the context for examining student assessment scores. Chapter 4 focuses on a second set of results from the 2000 reading assessment that includes the performance of special needs students who were permitted accommodations in the test administration. The chapter presents these results for the nation and selected subgroups of students, and compares them to the results presented in chapters 1 and 2.

1

Average Scale Score and Achievement Level Results for the Nation

This chapter presents the national results of the NAEP 2000 reading assessment at grade 4. Student performance is described by average scale scores on the NAEP reading composite scale and in terms of percentages of students who attained each of the three reading achievement levels: *Basic*, *Proficient*, and *Advanced*. Results of the NAEP 2000 reading assessment are compared with fourth-grade results from three previous assessments: 1992, 1994, and 1998. This comparison is possible because the assessments share a common set of reading tasks based on the current reading framework and because the populations of students were sampled and assessed using comparable procedures.

To illustrate the reading abilities demonstrated on the NAEP assessment, sample questions and actual student responses from the 2000 assessment are included in this chapter. Three sample questions are provided to show the kind of tasks students were asked to do and how they responded.

The concluding section of this chapter presents a map of selected item descriptions. The item map provides a more comprehensive portrait of student performance by placing item descriptions on the NAEP scale where the question was likely to be answered successfully by students. By mapping item descriptions in this way, the reading skills and abilities associated with different points on the scale are visually represented.

Chapter Focus

Are the nation's fourth-graders making progress in reading?

What reading abilities do they have?

Chapter Contents

Scale Score Results for the Nation

Achievement Level Results for the Nation

Sample Assessment Questions and Student Responses

Item Map

Average Scale Score Results

The results of the NAEP 2000 reading assessment at grade 4 show overall stability in student performance across the assessment years: the average reading scale score for 2000 was not significantly different from 1992, 1994, and 1998 results. Figure 1.1 presents the average reading scale scores of fourth-grade students attending both public and nonpublic schools for the four assessments between 1992 and 2000.

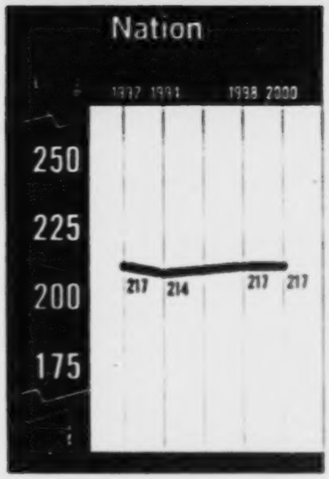
Another way to view students' reading ability is by looking at the scale scores attained by students across the performance distribution. The advantage of looking at the data this way is that it shows how fourth-graders with lower or higher reading ability performed in relation to the national average. In addition, the percentile data show whether trends in the national average score are reflected in scores across the performance distribution. Figure 1.2 presents the reading scale scores for fourth-grade students at the 10th, 25th, 50th, 75th, and 90th percentiles for the NAEP reading assessments from 1992–2000.

While the 2000 national average score of 217 was not significantly higher or lower in comparison to fourth-graders' average score in previous assessment years, stability across years is not reflected at all the percentiles. In fact, only at the 50th percentile—among students performing around the national average—have scores remained stable across assessments.

In 2000 the scores at both the 10th and 25th percentiles are higher in relation to 1994; however, at the 10th percentile the score in 2000 is lower than it was in 1992. A different trend is evident among higher-performing fourth-graders. At the 75th and 90th percentiles, scores have increased

Figure 1.1 Average fourth-grade reading scale scores: 1992–2000

Scale Score Results



SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

across assessment years, resulting in a 2000 score that is higher in relation to 1992. These data indicate that the lowest-performing students in 2000 were not performing as well as their counterparts in 1992; however, the results for the highest-performing students indicate an increase in performance between 1992 and 2000.

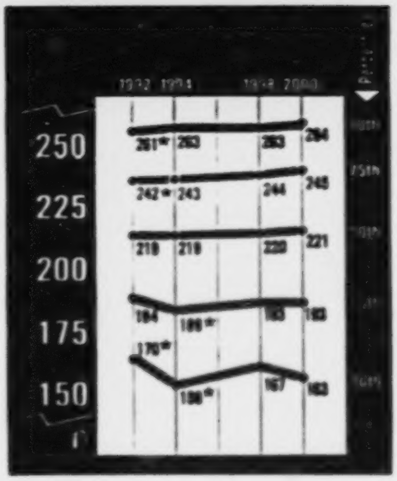
Achievement Level Results

The results of student performance are not only reported using scores on the NAEP reading scale, but also using reading achievement levels as authorized by the NAEP legislation and as adopted by the National Assessment Governing Board.[1] The achievement levels are performance standards adopted by the Board, based on the collective judgements of experts about what students should be expected to know and be able to do in terms of the NAEP reading framework. Viewing students' performance from this perspective provides some insight into the adequacy of students' knowledge and skills and the extent to which they achieved expected levels of performance.

The Board reviewed and adopted the recommended achievement levels in 1992, which were derived from the judgments of a broadly representative panel that included teachers, education specialists, and members of the general public. For each grade assessed, the Board has adopted three achievement levels: *Basic, Proficient,* and *Advanced*. For reporting purposes, the achievement level cut scores are placed

Figure 1.2 Fourth-grade reading scale score percentiles: 1992–2000

Performance Distribution



★ Significantly different from 2000.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

---

[1] The Improving America's Schools Act of 1994 (20 USC 9010) requires that the National Assessment Governing Board develop "appropriate student performance levels" for reporting NAEP results.

on the NAEP reading scale resulting in four ranges: the range below *Basic*, *Basic*, *Proficient*, and *Advanced*. Figure 1.3 presents specific descriptions of reading achievement for the levels at grade four.

The NAEP legislation requires that achievement levels be "used on a developmental basis until the Commissioner of Education Statistics determines . . . that such levels are reasonable, valid, and informative to the public." A discussion of the developmental status of achievement levels may be found in the introduction to this report.

Figure 1.3 **Descriptions of reading achievement for the levels at grade 4**

Reading Achievement Levels	
Basic **(208)**	Fourth-grade students performing at the *Basic* level should demonstrate an understanding of the overall meaning of what they read. When reading text appropriate for fourth-graders, they should be able to make relatively obvious connections between the text and their own experiences and extend the ideas in the text by making simple inferences. For example, when reading **literary text**, students should be able to tell what the story is generally about—providing details to support their understanding—and be able to connect aspects of the stories to their own experiences. When reading **informational text**, *Basic*-level fourth-graders should be able to tell what the selection is generally about or identify the purpose for reading it, provide details to support their understanding, and connect ideas from the text to their background knowledge and experiences.
Proficient **(238)**	Fourth-grade students performing at the *Proficient* level should be able to demonstrate an overall understanding of the text, providing inferential as well as literal information. When reading text appropriate to fourth grade, they should be able to extend the ideas in the text by making inferences, drawing conclusions, and making connections to their own experiences. The connection between the text and what the student infers should be clear. For example, when reading **literary text**, *Proficient*-level fourth-graders should be able to summarize the story, draw conclusions about the characters or plot, and recognize relationships such as cause and effect. When reading **informational text**, *Proficient*-level students should be able to summarize the information and identify the author's intent or purpose. They should be able to draw reasonable conclusions from the text, recognize relationships such as cause and effect or similarities and differences, and identify the meaning of the selection's key concepts.
Advanced **(268)**	Fourth-grade students performing at the *Advanced* level should be able to generalize about topics in the reading selection and demonstrate an awareness of how authors compose and use literary devices. When reading text appropriate to fourth grade, they should be able to judge text critically and, in general, give thorough answers that indicate careful thought. For example, when reading **literary text**, *Advanced*-level students should be able to make generalizations about the point of the story and extend its meaning by integrating personal experiences and other readings with the ideas suggested by the text. They should be able to identify literary devices such as figurative language. When reading **informational text**, *Advanced*-level fourth-graders should be able to explain the author's intent by using supporting material from the text. They should be able to make critical judgments of the form and content of the text and explain their judgments clearly.

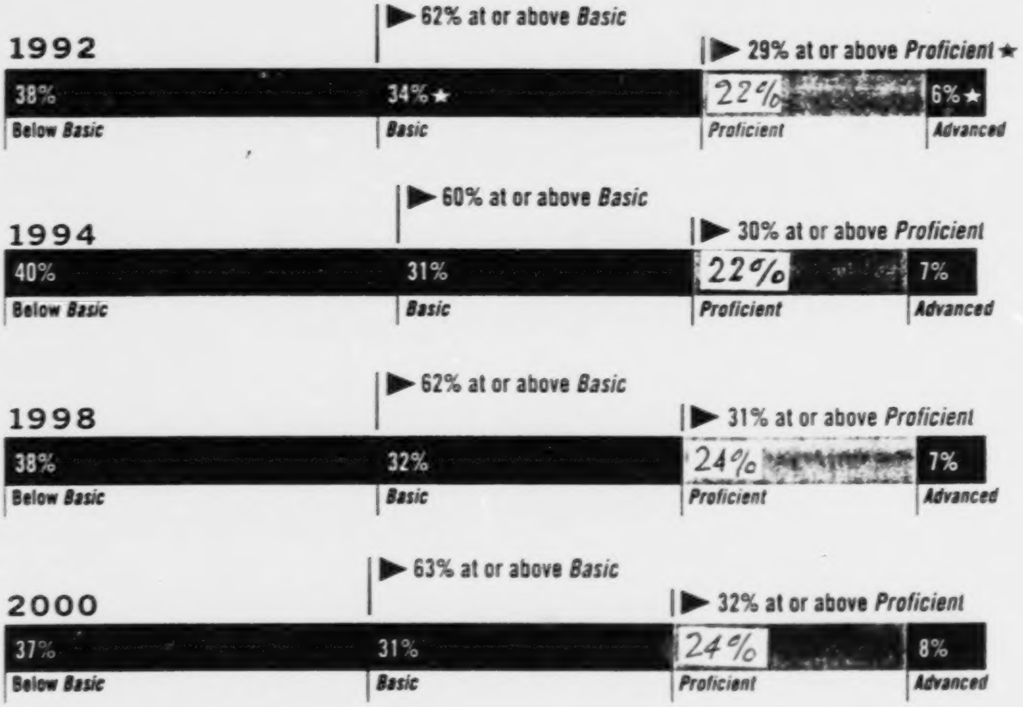
Achievement level results for the nation's fourth-grade students are presented in figure 1.4. Results are shown in two ways: the percentage of students within each achievement level interval, and the percentage of students at or above the *Basic* and at or above the *Proficient* achievement levels. In reading figure 1.4, it is necessary to keep in mind that the levels in the upper bars are cumulative; included among students who are considered to be at or above the *Basic* level are those who have also achieved the *Proficient* and *Advanced* levels of performance, and included among students who are considered to be at or above *Proficient* are those who have attained the *Advanced* level of performance.

As shown in figure 1.4, performance at

Figure 1.4

Achievement Level Results

Percentage of fourth-grade students at or above reading achievement levels and within each achievement level range: 1992–2000



★ Significantly different from 2000.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

or above the *Proficient* level—the achievement level identified by NAGB as the level that all students should reach—was achieved by 32 percent of students in 2000, while the highest level of performance, *Advanced*, was achieved by 8 percent of fourth-graders in 2000.

Viewing achievement level results across assessment years shows that in 2000 a higher percentage of students were at or above the *Proficient* level and at the *Advanced* level in comparison to 1992. Corresponding to the increase at or above the *Proficient* level, figure 1.4 shows that a lower percentage of students fell within the *Basic* level in 2000 than in 1992.

Sample Assessment Questions and Student Responses

The following pages present sample questions and student responses that portray student performance on the 2000 NAEP reading assessment. Three questions, which were asked about an informative article, were selected to exemplify the range of reading abilities demonstrated by students. The full text of the informative article, as well as additional sample questions, can be found in appendix B.

The three questions to follow include a multiple-choice, a short constructed-response, and an extended constructed-response question. The correct oval is filled in for the multiple-choice question. For constructed-response questions, a summary of the scoring criteria used to rate students' responses is provided. Actual student responses have been reproduced from assessment test booklets to illustrate representative answers. The rating assigned to each sample response is indicated.

The tables in this section present two types of percentages for each sample question: the overall percentage of students who answered successfully, and the percentage of students who answered successfully within a specific score range on the NAEP reading composite scale. The score ranges correspond to the three achievement level intervals—*Basic*, *Proficient*, and *Advanced*. It should be noted that the overall percentage of students shown in these tables includes students in the range below *Basic*, as well as students whose performance fell within the three achievement level ranges.

Informative Article:

A Brick to Cuddle Up To is an informational passage describing what colonists did to keep warm during the winter months. The author includes descriptions of how colonial homes were heated, how colonists kept warm at night, how they heated water for baths, as well as providing details that depicted differences between contemporary and colonial American life. (See appendix B for full text of passage.)

Sample Question 1:

In writing this article, the author mostly made use of

Ⓐ broad ideas

● specific details

Ⓒ important questions

Ⓓ interesting characters

Reading Purpose:
To Gain Information

Reading Stance:
Critical Stance

Table 1.1: Sample question 1 results (multiple-choice)

Overall percentage correct and percentages correct within each achievement level range: 2000

	Percentage correct within achievement level intervals		
Overall percentage correct[†]	***Basic*** **208–237***	***Proficient*** **238–267***	***Advanced*** **268 and above***
66	72	79	[illegible]

[†] Includes fourth-grade students who were below the *Basic* level.

*NAEP Reading composite scale range.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Sample Question 2:

Reading Purpose:
To Gain Information

Reading Stance:
Developing an Interpretation

Do you think "A Brick to Cuddle Up To" is a good title for this article? Using information from the article, tell why or why not.

Responses to this question were scored according to a three-level rubric as Unsatisfactory, Partial, or Complete

Table 1.2: Sample question 2 results (short constructed-response)

Overall percentage "Complete" and percentages "Complete" within each achievement level range: 2000

	Percentage "Complete" within achievement level intervals		
Overall percentage "Complete"†	***Basic* 208–237***	***Proficient* 238–267***	***Advanced* 268 and above***
37	38	57	76

† Includes fourth-grade students who were below the *Basic* level.

*NAEP Reading composite scale range.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Responses scored "Complete" demonstrated understanding of how the title relates to the central theme by indicating that the article described methods of keeping warm during winter in colonial times.

Sample "Complete" Response:

Do you think "A Brick to Cuddle Up To" is a good title for this article? Using information from the article, tell why or why not.

Yes I do think it is a good title. It is a good title because the article tells about how colonists kept warm in the winter & how they used heated bricks to keep warmen

Sample Question 3:

Pretend that you are an early American colonist. Describe at least three activities you might do during a cold winter evening. Be specific. Use details from the article to help you write your description.

Reading Purpose:
To Gain Information

Reading Stance:
Developing an Interpretation

Responses to this question were scored according to a four-level rubric as Unsatisfactory, Partial, Essential, or Extensive

Table 1.3: Sample question 3 results (extended constructed-response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement level range: 2000

	Percentage "Essential" or better within achievement level intervals		
Overall percentage "Essential" or better[1]	***Basic*** **208–237***	***Proficient*** **238–267***	***Advanced*** **268 and above***
18	15	29	40

[1] Includes fourth-grade students who were below the *Basic* level.

*NAEP Reading composite scale range.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Responses scored "Essential" demonstrated comprehension of colonial life as portrayed in the article by providing three activities, some of which are related to the need to stay warm.

Sample "Essential" Response:

Pretend that you are an early American colonist. Describe at least three activities you might do during a cold winter evening. Be specific. Use details from the article to help you write your description.

I would sit by the Fireplace
and read, or do needlework
or stir a kettle or corn
pudding, or check the baking
bread.

Responses scored "Extensive" demonstrated comprehension of the central theme of the article. Of the activities provided, at least three focus on the need to stay warm.

Sample "Extensive" Response:

Pretend that you are an early American colonist. Describe at least three activities you might do during a cold winter evening. Be specific. Use details from the article to help you write your description.

On a cold winter evening I would probly get out a game and play it with my family. We would have a nice hot snack and drink to keep them warm. We would all sit around the fireplace, maybe tell old storys when we were little. We could tell ghost stories. When they went to bed I would put the warm brick under their bed While they were asleep. I would Knit a warm blanket for church or trips..

Map of Selected Item Descriptions

The reading performance of fourth-graders and aspects of reading comprehension can be illustrated by a map that positions item descriptions along the NAEP reading composite scale where items are likely to be answered successfully by students.[2] The descriptions used on the map focus on the reading skill or ability needed to answer the question. For multiple-choice questions, the description indicates the comprehension demonstrated by selection of the correct response option; for constructed-response questions, the description takes into account the degree of comprehension demonstrated as specified by the different levels of scoring criteria for that question. An examination of the descriptions provides some insight into the range of comprehension processes demonstrated by fourth-grade students at different score points on the NAEP scale.

For each question indicated on the map, students who scored above the scale point had a higher probability of successfully answering the question, and students who scored below the scale point had a lower probability of successfully answering the question. The map identifies where individual comprehension questions were answered successfully by at least 65 percent of the students for constructed-response questions, or by at least 74 percent of the students for multiple-choice questions.[3] For example, a multiple-choice question that asks students to identify the major topic of an article maps at 224 on the scale. This means that fourth-grade students with an average scale score of 224 or more have at least a 74 percent chance of answering this question correctly. In other words, at least 74 out of every 100 students who score at or above 224 answered this question correctly. Although students scoring above the scale point had a higher probability of successfully answering the question, it does not mean that every student at or above 224 always answers this question correctly, nor does it mean that students below 224 always answered the question incorrectly. The item maps are a useful indicator of higher or lower probability of successfully answering the question depending on students' overall ability as measured by the NAEP scale.

In considering the information provided by the item maps, it is important to recognize that these descriptions are based on comprehension questions that were answered about specific passages. It is possible that questions intended to assess the same aspect of comprehension, when asked about different passages, would map at different points on the scale. In fact, one NAEP study found that even identically worded questions function differently (i.e., easier or harder) when associated with different passages, suggesting that the difficulty of a question resides not only in the question itself, but also in the interaction of the question with a particular passage.[4]

---

[2] Details on the procedures used to develop item maps will be provided in the forthcoming *NAEP 2000 Technical Report*.

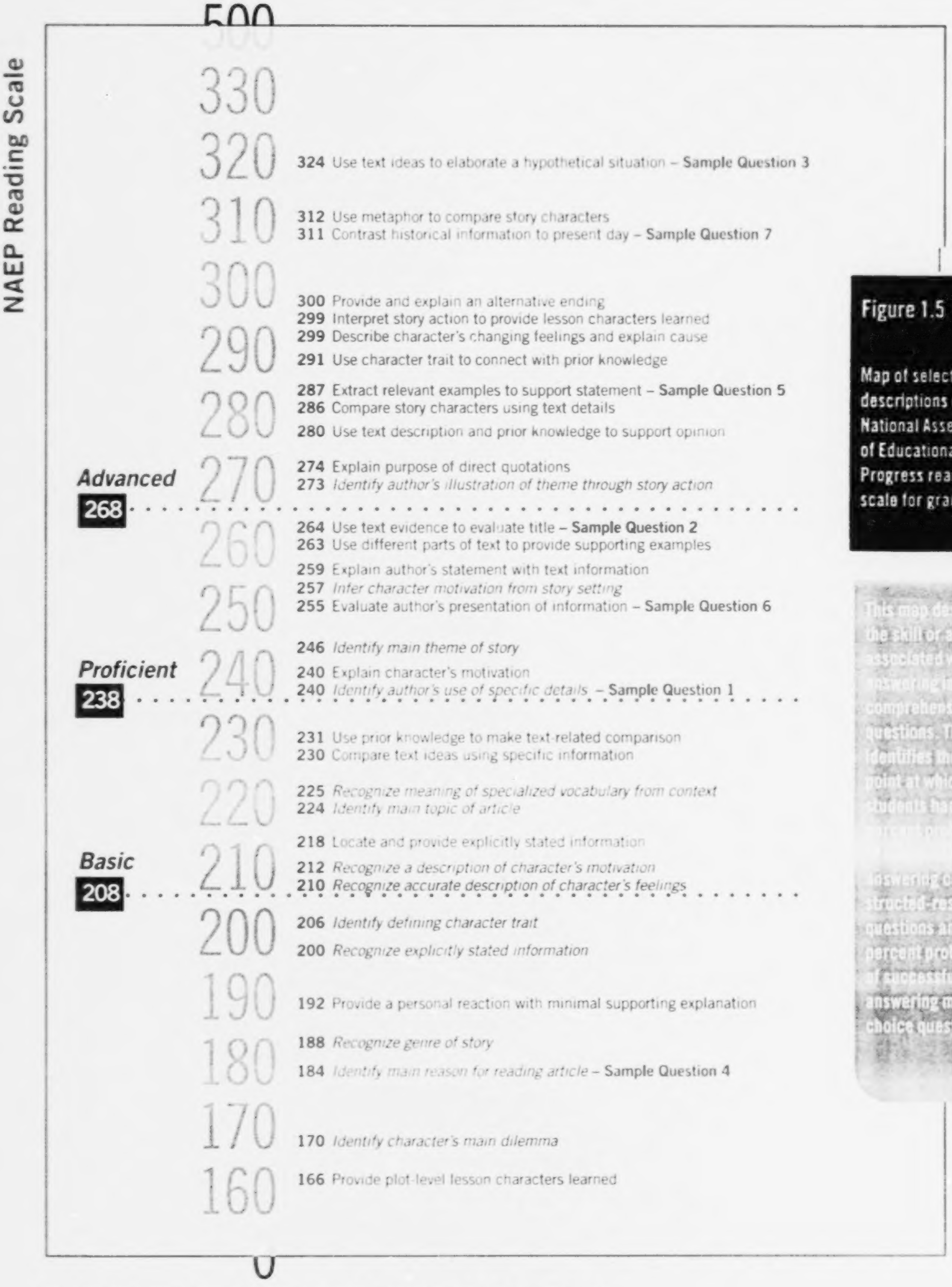
[3] The probability convention is set higher, at 74 percent, for multiple-choice questions to correct for the possibility of answering correctly by guessing.

[4] Campbell, J.R. & Donahue, P.L. (1997). *Students selecting stories: The effects of choice in reading assessment.* Washington, DC: National Center for Education Statistics.

Figure 1.5

Map of selected item descriptions on the National Assessment of Educational Progress reading scale for grade 4

This map describes the skill or ability associated with answering individual comprehension questions. The map identifies the score point at which students had a 65 percent probability of successfully answering constructed-response questions and a 74 percent probability of successfully answering multiple-choice questions.



324 Use text ideas to elaborate a hypothetical situation – **Sample Question 3**
312 Use metaphor to compare story characters
311 Contrast historical information to present day – **Sample Question 7**
300 Provide and explain an alternative ending
299 Interpret story action to provide lesson characters learned
299 Describe character's changing feelings and explain cause
291 Use character trait to connect with prior knowledge
287 Extract relevant examples to support statement – **Sample Question 5**
286 Compare story characters using text details
280 Use text description and prior knowledge to support opinion
274 Explain purpose of direct quotations
273 *Identify author's illustration of theme through story action*

Advanced **268**

264 Use text evidence to evaluate title – **Sample Question 2**
263 Use different parts of text to provide supporting examples
259 Explain author's statement with text information
257 *Infer character motivation from story setting*
255 Evaluate author's presentation of information – **Sample Question 6**
246 *Identify main theme of story*
240 Explain character's motivation
240 *Identify author's use of specific details* – **Sample Question 1**

Proficient **238**

231 Use prior knowledge to make text-related comparison
230 Compare text ideas using specific information
225 *Recognize meaning of specialized vocabulary from context*
224 *Identify main topic of article*
218 Locate and provide explicitly stated information
212 *Recognize a description of character's motivation*
210 *Recognize accurate description of character's feelings*

Basic **208**

206 *Identify defining character trait*
200 *Recognize explicitly stated information*
192 Provide a personal reaction with minimal supporting explanation
188 *Recognize genre of story*
184 *Identify main reason for reading article* – **Sample Question 4**
170 *Identify character's main dilemma*
166 Provide plot-level lesson characters learned

NOTE: Regular type denotes a constructed-response question. Italic type denotes a multiple-choice question.

Each grade 4 reading question was mapped onto the NAEP 0–500 reading scale. The position of the question on the scale represents the scale score attained by students who had a 65 percent probability of successfully answering a constructed-response question or a 74 percent probability of correctly answering a four-option, multiple-choice question. Only selected questions are presented. Scale score ranges for reading achievement levels are referenced on the map.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Summary

This chapter presented overall results from the NAEP reading assessment for fourth-grade students. Results examined included the national average score, the percentile scores across the performance distribution, and attainment of the reading achievement levels.

The following figure displays the major findings presented in this chapter. In each line of the display, the average reading score, the percentile score, or the percentage of students at or above achievement levels is compared to that in the first assessment year under the current reading framework. An arrow pointing upward (↑) indicates a significant increase, a horizontal arrow (→) indicates no significant change, and an arrow pointing downward (↓) indicates a significant decrease. For example, the first section of the display shows that the national average for fourth-grade students has not changed significantly since 1992, whereas the first line in the next section indicates that the score at the 90th percentile has increased since 1992.

Figure 1.6

Chapter Summary

Average scale score, percentile scores, and achievement level results: 2000 compared to 1992

National Average

→ Scale score since 1992

Percentile Scores

↑ 90th percentile score since 1992
↑ 75th percentile score since 1992
→ 50th percentile score since 1992
→ 25th percentile score since 1992
↓ 10th percentile score since 1992

Achievement Level Results

↑ Percentage at *Advanced* since 1992
↑ Percentage at or above *Proficient* since 1992
→ Percentage at or above *Basic* since 1992

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 2000 Reading Assessments.

2

Average Scale Score and Achievement Level Results for Selected Subgroups of 4th-Grade Students

Implicit in the call for high standards is the goal that all children regardless of background or where they attend school should be held to the same expectations and given the same opportunity for achievement. Within this larger context of reform, the performance results in this chapter may prove useful as one indicator of progress toward the nation's educational goals. Presented in this chapter are results for six subgroups of students with different demographic characteristics. Results are reported by: gender, race/ethnicity, region of the country, school's type of location, eligibility for the free/reduced-price lunch program, and type of school. For all subgroups except type of location and free/reduced-price lunch eligibility, results of the 2000 assessment can be compared to the results of previous assessments under the current framework, 1992, 1994, and 1998.

Results are presented both in terms of students' scores on the NAEP reading composite scale and in terms of the percentages of students attaining achievement levels. Scale score results represent students' actual performance on the assessment, and the achievement level results view that performance in relation to set expectations of what students should know and be able to do.

Chapter Focus

Are selected subgroups of students making progress in reading?

Chapter Contents

Gender

Race/Ethnicity

Trends in Scale Score Differences

Region of the Country

Type of Location

Eligibility for the Free/Reduced-Price Lunch Program

Type of School

The reading achievement levels—*Basic, Proficient,* and *Advanced*—used to report NAEP results were established by the National Assessment Governing Board (NAGB) in 1992 for the content framework that provides the basis for the reading assessments. Descriptions of abilities associated with the three levels are presented in figure 1.3 in chapter 1 of this report.

The NAEP legislation (The Improving America's Schools Act of 1994) requires that achievement levels be "used on a developmental basis until the Commissioner of Education Statistics determines . . . that such levels are reasonable, valid, and informative to the public." A discussion of the developmental status of the achievement levels is included in the introduction to this report.

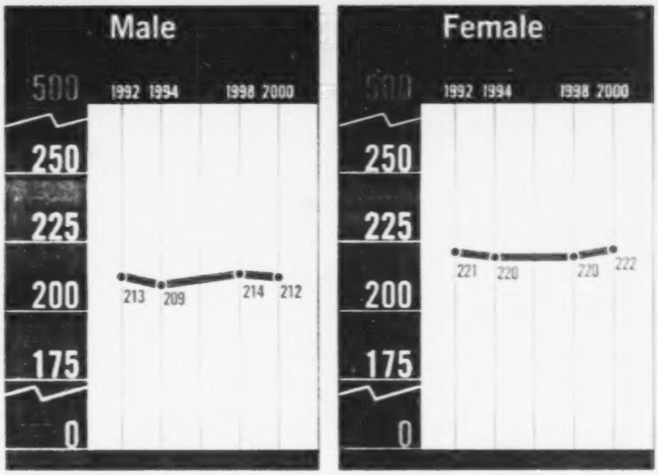
Differences reported in this chapter between demographic subgroups for the 2000 assessment and between previous assessments are based on statistical tests that consider both the magnitude of the difference between the group average scores or percentages and the standard errors of those statistics. Differences between groups and between assessment years are discussed only if they have been determined to be statistically significant.

Gender

Figure 2.1 presents average reading scale scores for male and female students across assessment years. In the 2000 assessment, no significant increase or decrease was observed for either male or female fourth-grade students. The increase achieved by male students between 1994 and 1998 did not continue or lead to a higher score in 2000. The average score for female fourth-graders has remained relatively consistent across assessment years. In 2000, results show female students continuing to outperform their male counterparts.

Figure 2.1 **Average fourth-grade reading scale scores for male and female students: 1992–2000**

Scale Score Results by Gender



SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

The percentages of male and female students at or above the reading achievement levels and within each achievement level range are presented in figure 2.2. Among male students, apparent changes in the percentages at or above any of the achievement levels across assessment years were not statistically significant. In 2000, however, the percentage of females at or above the *Proficient* level (36 percent) was higher in comparison to 1992 (32 percent). The achievement level results are consistent with the scale score results, in that they also show female students outperforming male students. Comparison of male and female performance in 2000 shows higher percentages of female fourth-graders at or above *Basic*, at or above *Proficient*, and at the *Advanced* level.

Figure 2.2
Achievement Level Results by Gender

Percentages of fourth-grade male and female students at or above reading achievement levels and within each achievement level range: 1992–2000

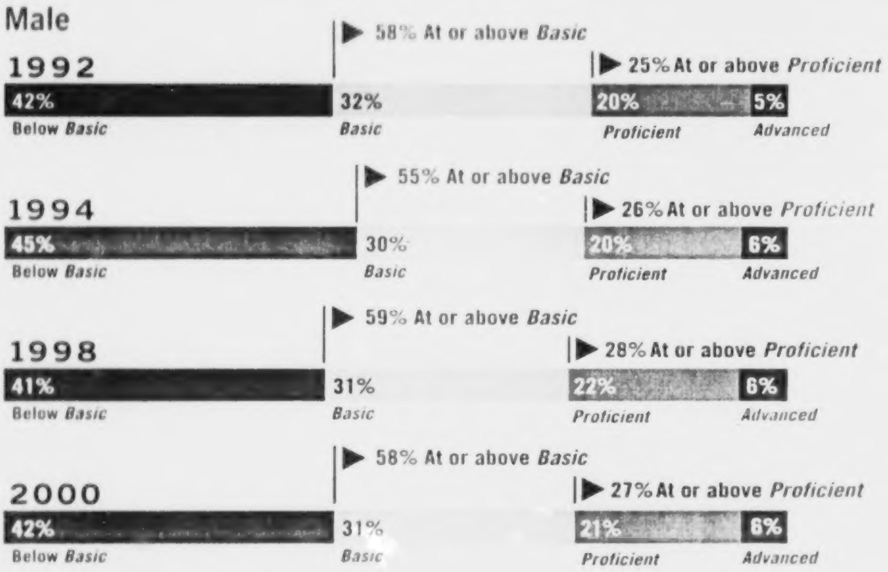
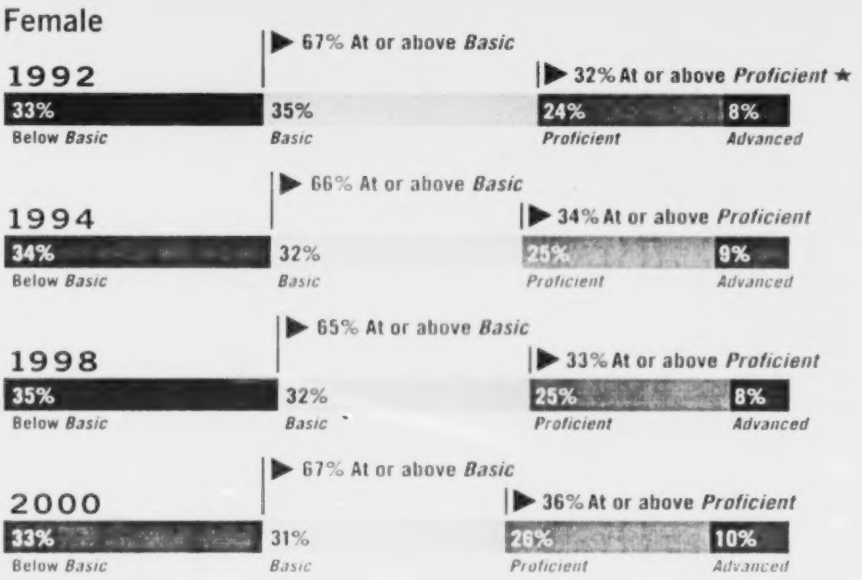


Figure continues on the next page.

Figure 2.2

Achievement Level Results by Gender (continued)

Percentages of fourth-grade male and female students at or above reading achievement levels and within each achievement level range: 1992–2000



★ Significantly different from 2000.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Race/Ethnicity

The background questionnaire administered with the 2000 NAEP reading assessment asked students to indicate the racial/ethnic subgroup that best described them. In the 1992, 1998, and 2000 reading assessments the mutually exclusive subgroup categories were: White, Black, Hispanic, Asian/Pacific Islander, and American Indian (including Alaskan native). In 1994, a similar background question was asked, the only difference being that Asian and Pacific Islander were two separate response categories. To analyze changes in performance across assessment years, the separate response categories used in 1994 were combined into a single category: Asian/Pacific Islander, as used in the other assessment years.

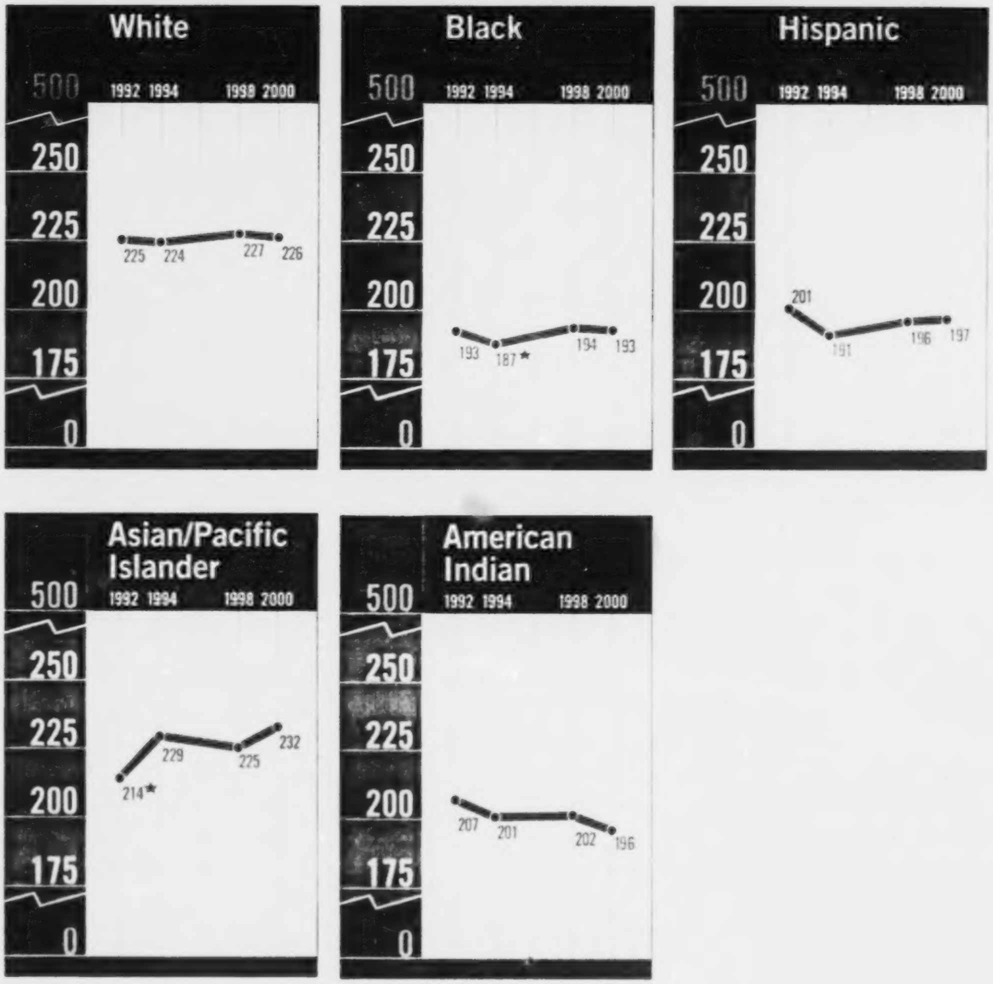
Across assessment years, some changes are evident in the population of fourth-graders sampled by NAEP. The population of Hispanic fourth-grade students in the NAEP sample has increased from 9 percent in 1992 to 15 percent in 2000, while the population of white fourth-grade students has decreased from 71 percent in 1992 to 66 percent in 2000. The across-year percentages for all racial/ethnic subgroups can be found in table C.9 in appendix C.

The average reading scale scores for students by racial/ethnic subgroup across assessment years are presented in figure 2.3. Of the five racial/ethnic subgroups, only Asian/Pacific Islander students showed overall gains since 1992. The 2000 average score of students in each of the other subgroups was similar to or did not differ significantly from the 1992 average score.

Although black and Hispanic students' average scores had declined in 1994, slight rebounds since that time are evident for both groups. For black students only, however, this resulted in a 2000 average score that was significantly higher than that in 1994. Any apparent differences between

Figure 2.3 **Scale Score Results by Race/Ethnicity**

Average fourth-grade reading scale scores by race/ethnicity: 1992–2000



★ Significantly different from 2000.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

years in the average scores of white or American Indian students were not statistically significant. Comparisons of subgroup performance in 2000 show white and Asian/Pacific Islander students outperforming their black, Hispanic, and American Indian peers.

Achievement level results for racial/ethnic subgroups are presented in figure 2.4. Consistent with scale score results, achievement level results show that the only group of students with significant gains was Asian/Pacific Islander students. A higher percentage of Asian/Pacific Islander students were at or above the *Proficient* level in 2000 than in 1992. For other racial/ethnic subgroups, any apparent differences from past assessments observed in the percentages of students at or above the *Basic* and *Proficient* achievement levels were not statistically significant. Comparisons of achievement level results between subgroups in 2000 show higher attainment for white and Asian/Pacific Islander students. The percentages of white and Asian/Pacific Islander students at or above *Proficient* and at or above *Basic* were higher than the percentages for all other racial/ethnic groups.

Figure 2.4
Achievement Level Results by Race/Ethnicity

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by race/ethnicity: 1992–2000

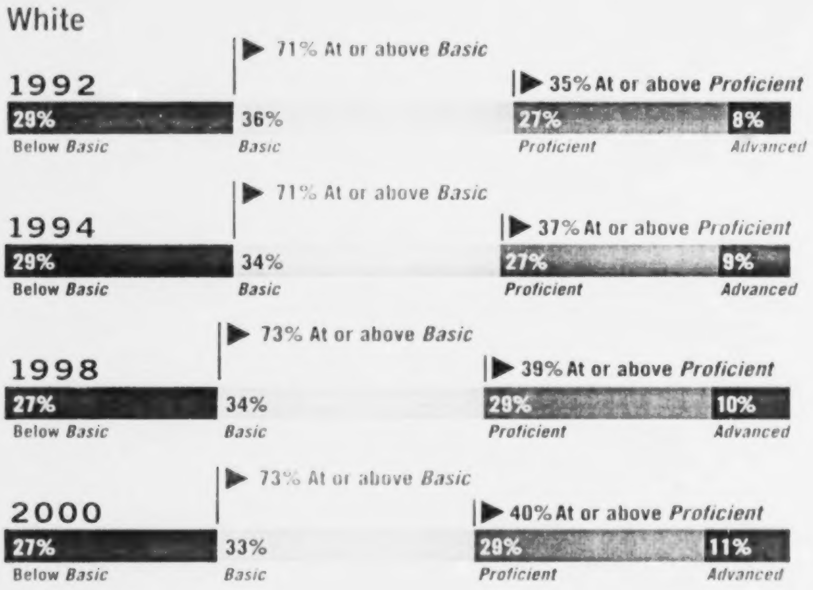


Figure 2.4
Achievement Level Results by Race/Ethnicity (continued)

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by race/ethnicity: 1992–2000

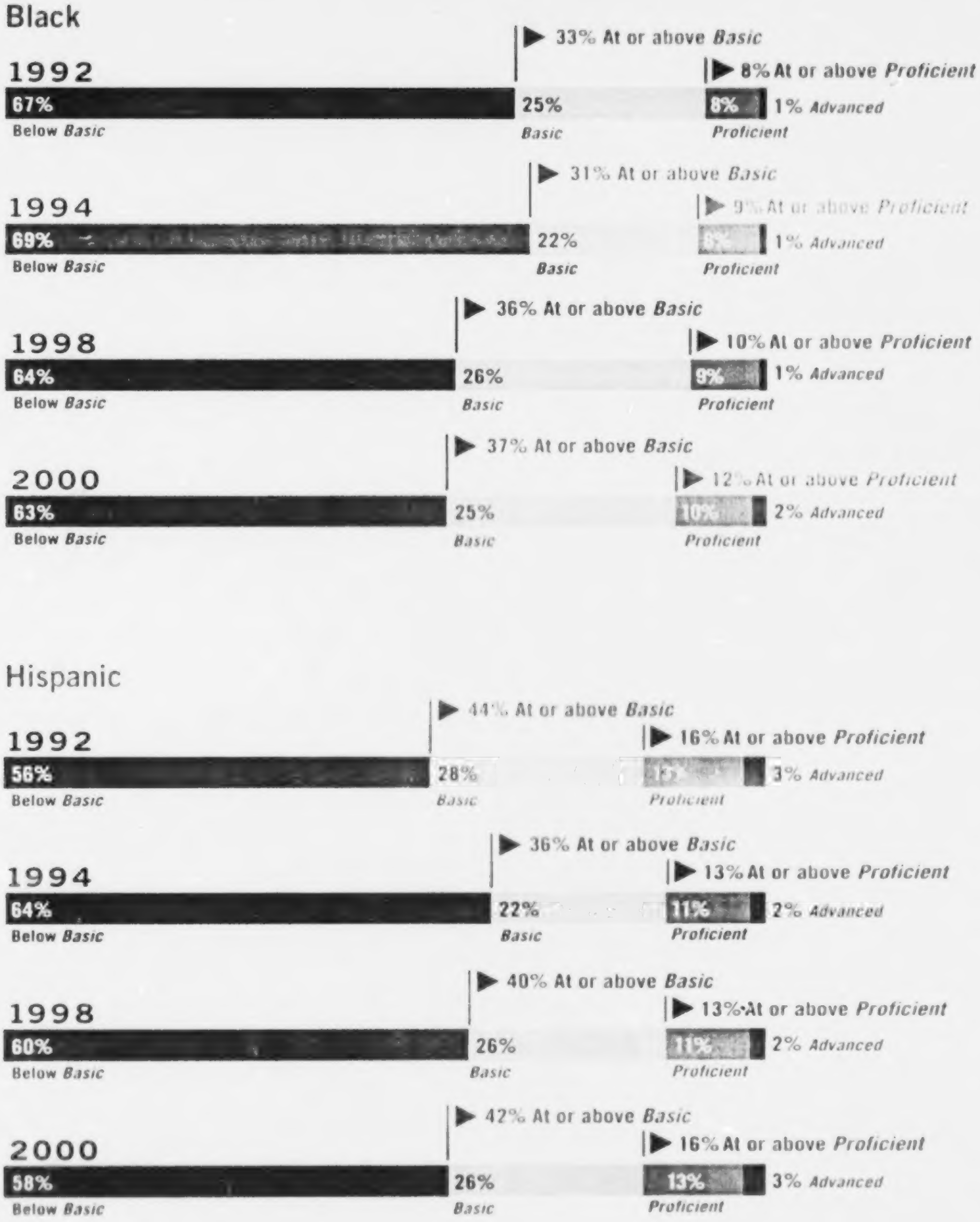
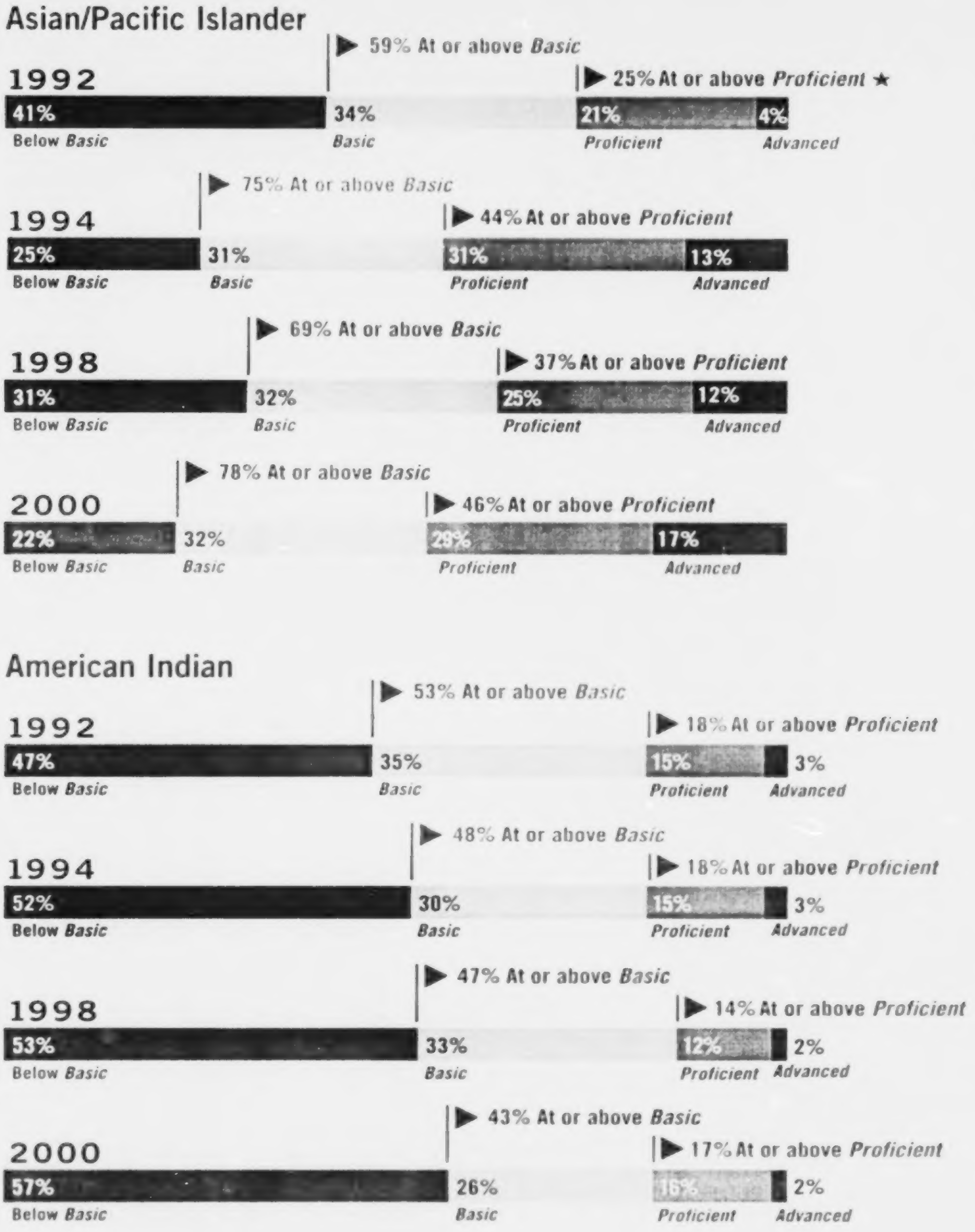


Figure continues on the next page

Figure 2.4
Achievement Level Results by Race/ Ethnicity (continued)

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by race/ethnicity: 1992–2000



★ Significantly different from 2000.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Trends in Scale Score Differences Between Selected Subgroups

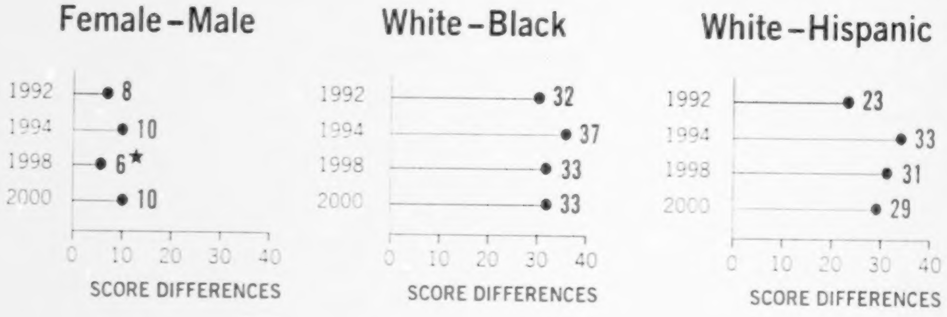
As with previous NAEP reading assessments, results from the 2000 NAEP continue to show performance differences between racial/ethnic subgroups and between male and female students. Such differences between the average assessment scores of student groups should be interpreted cautiously. The average score of a selected group does not represent the entire range of performance within that group. Furthermore, differences between groups of students can not be attributed solely to group identification. A complex array of educational and social factors interacts to affect average student performance.

Scale score differences between selected subgroups of students from 1992 to 2000 are displayed in figure 2.5. In the 2000 assessment, white students outperformed their black and Hispanic peers by 33 and 29 scale points, respectively. Female students outperformed male students by 10 scale points.

Little change is seen since 1992 in the magnitude of the gap between these groups of students. The gap between white and black, and between white and Hispanic students, has varied somewhat across the assessment years, but none of the changes were statistically significant. The gap between male and female students has also fluctuated slightly; however, the 4-point increase in score differences between males and females between 1998 and 2000 represents a statistically significant widening of the gender gap.

Figure 2.5
Scale Score Differences Between Selected Subgroups

Differences in average fourth-grade reading scale scores by gender and race/ethnicity: 1992–2000



★ Significantly different from 2000.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

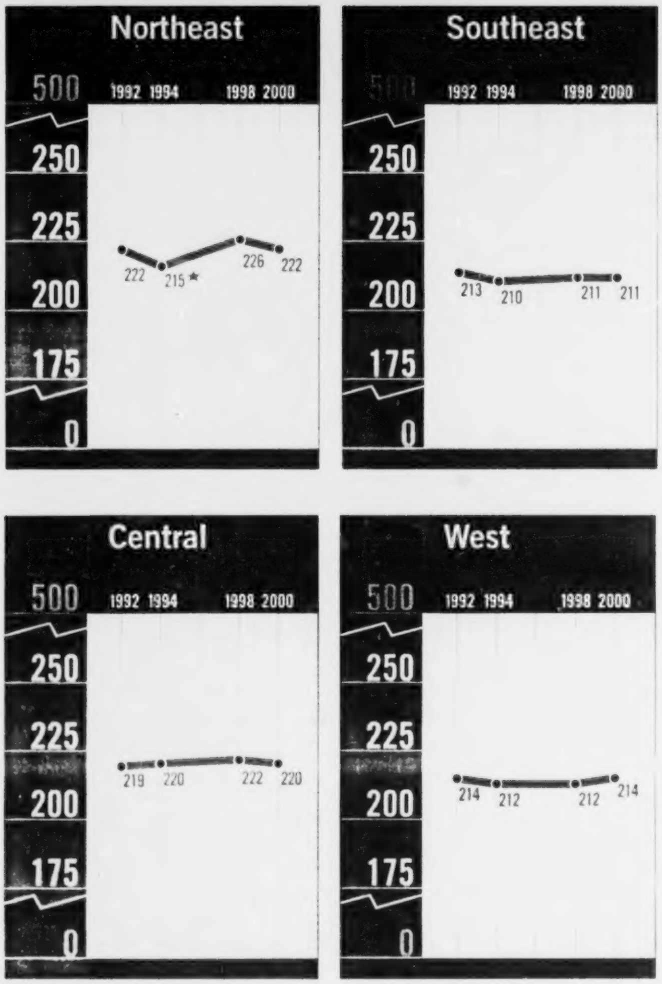
Region of the Country

This section examines results for four regions of the country: Northeast, Southeast, Central, and West. A listing of the states that are within these regions is provided in appendix A.

Figure 2.6 presents scale score results by region. Most of the gains made between 1994 and 1998 by fourth-graders in the Northeast were maintained; their average score in 2000 was higher than in 1994, but not significantly different from 1992. For

Figure 2.6
Scale Score Results by Region

Average fourth-grade reading scale scores by region of the country: 1992–2000



★ Significantly different from 2000

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

the other regions, student performance in 2000 varied only slightly from previous assessment years, and none of the apparent variations were statistically significant.

Comparisons between the regions in the 2000 assessment show fourth-graders in the Northeast and Central regions outperforming their peers in the Southeast and the West.

As shown in figure 2.7, achievement level results were stable for all the regions. No increases or decreases were observed in the percentages of students at or above any of the achievement levels. While the percentage of students in the Northeast who were at or above the *Basic* level had increased between 1994 and 1998, the percentage at or above *Basic* in 2000 was not significantly higher than in past assessments.

Comparisons of achievement level results between the regions show higher attainment by fourth-grade students in the Northeast and Central regions. The Northeast region had a higher percentage of students at the *Proficient* level and at the *Advanced* level than the Southeast. The percentages of Northeast students at or above *Proficient* and at or above the *Basic* level were higher than both the Southeast and the West. The Central region had higher percentages of students at *Proficient*, as well as at or above *Proficient* and at or above *Basic* than the Southeast.

Figure 2.7
Achievement Level Results by Region

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by region of the country: 1992–2000

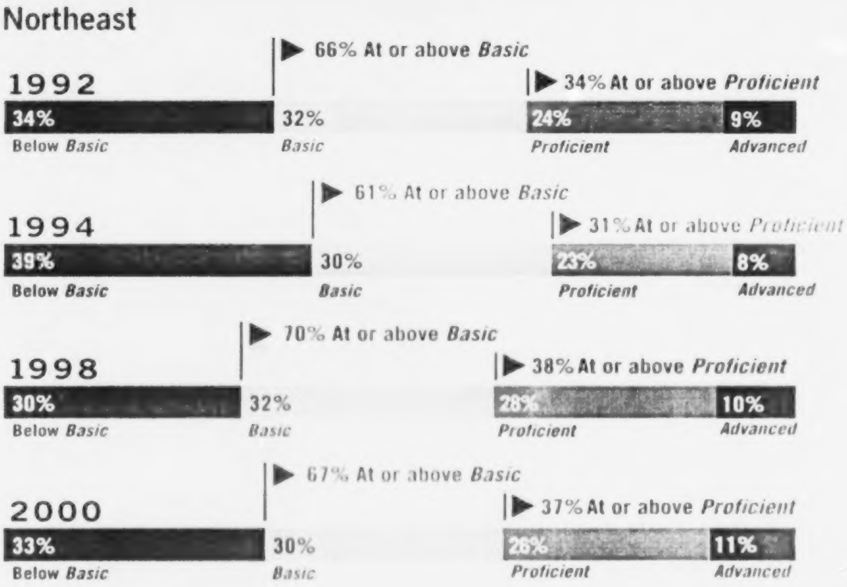


Figure continues on the next page.

Figure 2.7
Achievement Level Results by Region (continued)

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by region of the country: 1992–2000

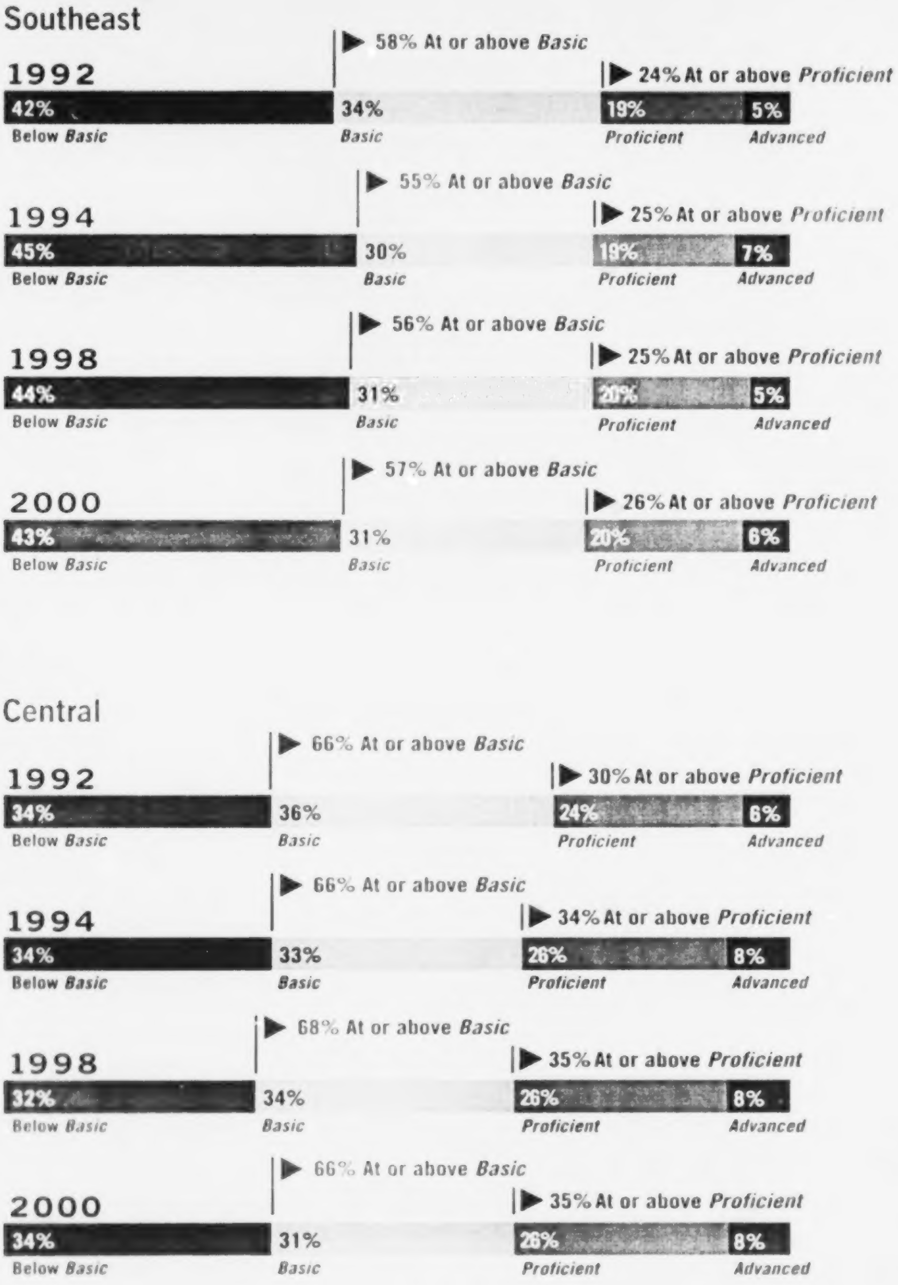
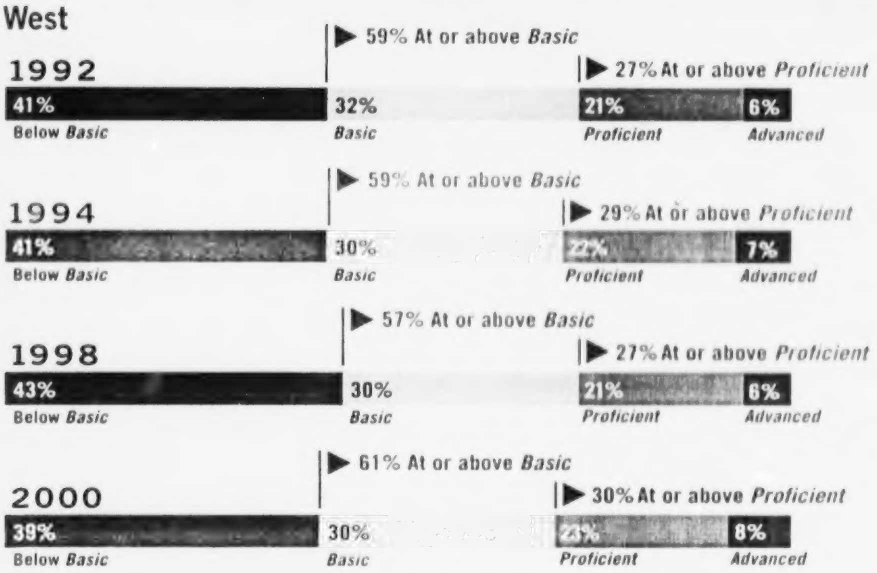


Figure 2.7
Achievement Level Results by Region (continued)

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by region of the country: 1992–2000



NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Type of Location

The schools of students who participate in the NAEP assessments are classified according to type of location. Based on Census Bureau definitions of metropolitan statistical areas, population size, and density, the three mutually exclusive categories are: central city, urban fringe/large town, and rural/small town. Due to a change in how locations were identified in 2000, comparisons to previous assessment years based on type of location are not possible. An explanation of this change can be found in appendix A.

Table 2.1
Scale Score Results by Type of Location

Average fourth-grade reading scale scores by school's type of location: 2000

	Central city	Urban fringe/large town	Rural/small town
2000	209	222	218

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

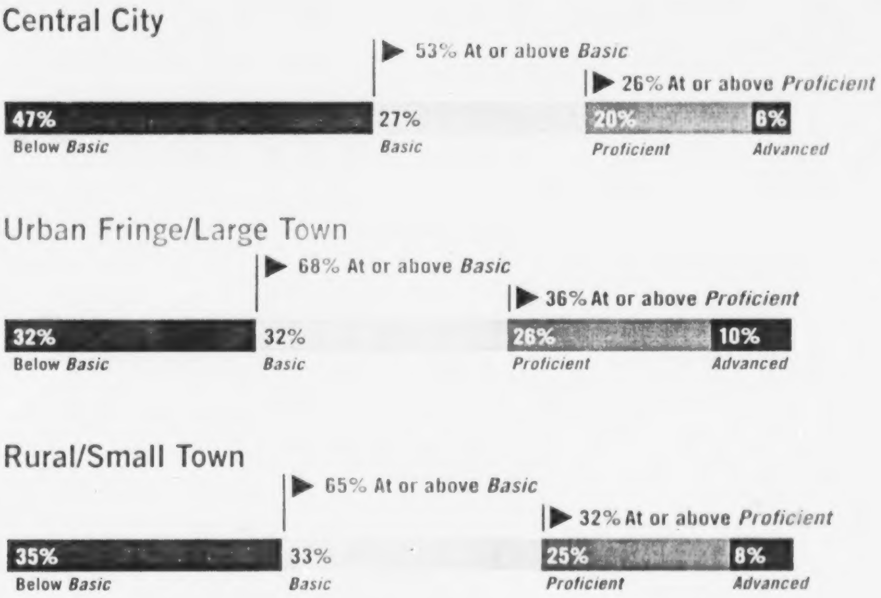
BEST COPY AVAILABLE

Consistent with past performance patterns when comparing location types, students in central city schools had a lower average reading score than their peers who attend schools in other types of location.

Figure 2.8 presents achievement level results by type of location. Comparisons of achievement level results between locations show lower percentages of central city students at or above the *Proficient* level and at or above the *Basic* level of performance than their peers in urban/fringe and rural locations. The slight differences between locations in the percentage of students who attained the *Advanced* level were not statistically significant.

Figure 2.8
Achievement Level Results by Type of Location

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by school's type of location: 2000



NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

BEST COPY AVAILABLE

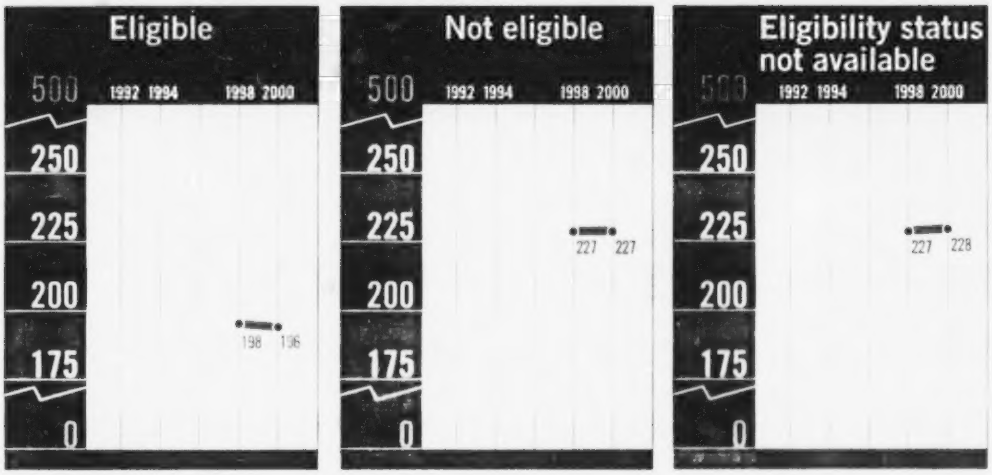
Eligibility for the Free/Reduced-Price Lunch Program

Funded by the U.S. Department of Agriculture (USDA) as part of the National School Lunch Program (NSLP), the free/reduced-price lunch program is available to public schools, nonprofit private schools, and residential childcare institutions. Eligibility for free or reduced-price lunch is determined through the USDA Income Eligibility Guidelines whereby students from families near or below the poverty line can be assured of having a wholesome meal at school.[1]

While NAEP first collected information on student eligibility for the federally funded NSLP in 1996, it was not until 1998 that this indicator of poverty was reported for students participating in the NAEP reading assessment. Thus comparisons can only be made between the two most recent assessments. As shown in figure 2.9, there was no change in average scores between the two assessment years for non-eligible students and the apparent decrease for eligible students was not statistically significant. In 2000 the score for eligible students remained significantly lower than that for students not eligible to receive a free or reduced-price lunch.

Figure 2.9
Scale Score Results by Free/Reduced-Price Lunch Eligibility

Average fourth-grade reading scale scores by student eligibility for the free/reduced-price lunch program: 1998–2000



SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

---

[1] U.S. General Services Administration. (1999). *Catalogue of federal domestic assistance.* Washington, DC: Executive Office of the President, Office of Management and Budget.

As shown in figure 2.10, no change was observed since 1998 in the percentages of eligible and noneligible students attaining specific achievement levels. As with average scale score results, achievement level results also show lower performance among students eligible for the program. Fourteen percent of eligible students performed at or above the *Proficient* achievement level in comparison to 41 percent of noneligible students. Among fourth-graders who were eligible for the program, 60 percent were below the *Basic* level.

Figure 2.10
Achievement Level Results by Free/Reduced-Price Lunch Eligibility

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by student eligibility for the free/reduced-price lunch program: 1998–2000

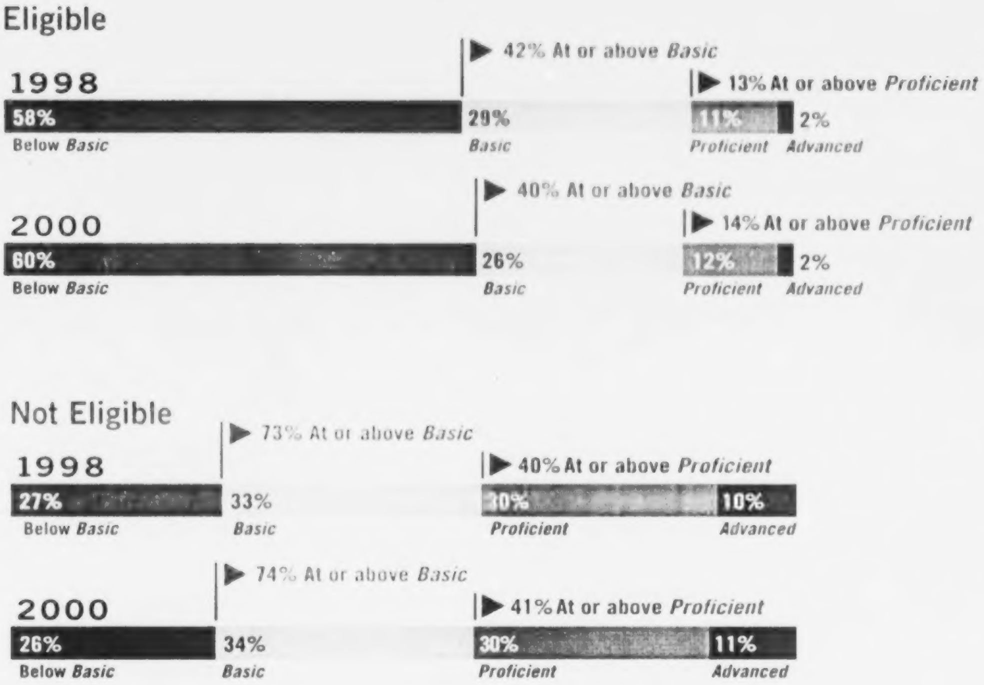
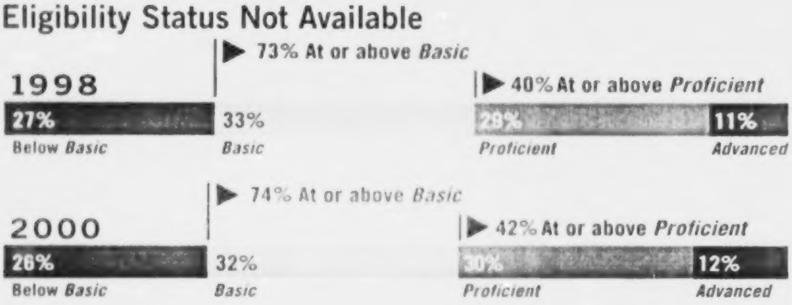


Figure 2.10
Achievement Level Results by Free/Reduced-Price Lunch Eligibility (continued)

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by student eligibility for the free/reduced-price lunch program: 1998–2000



NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Type of School

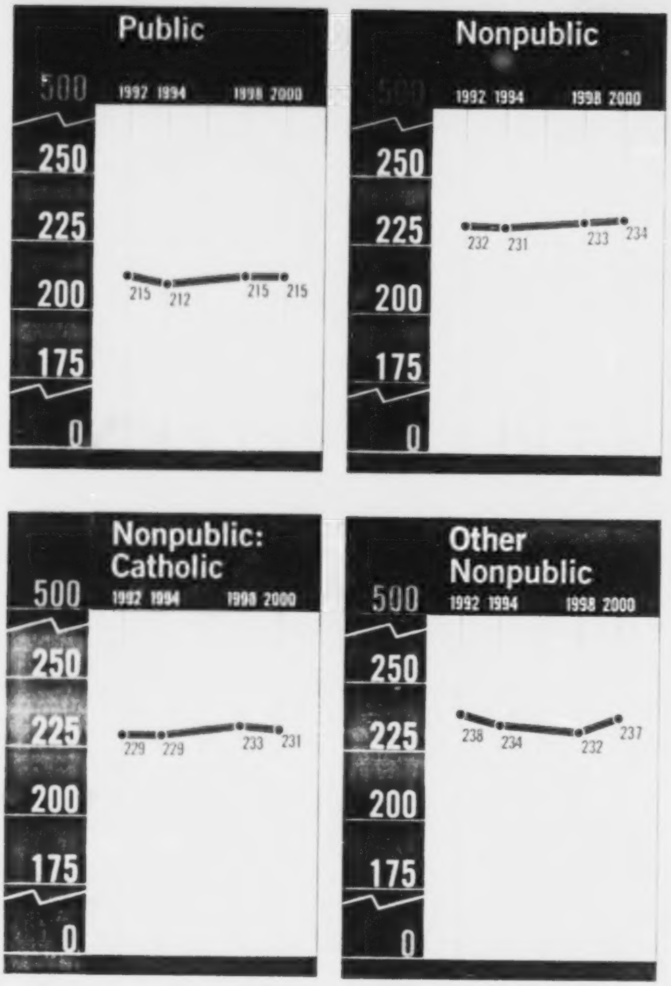
The schools of students who take the NAEP reading assessment are first classified overall as either public or nonpublic. A further distinction is then made within the nonpublic classification between schools that are Catholic and other nonpublic schools. Differences in performance between public and nonpublic schools surveyed and reported by the NAEP reading assessment have consistently shown that students attending the various types of nonpublic schools outperform their public school peers. Despite this consistent pattern in performance results, readers are cautioned against making assumptions about the comparative quality of instruction in public and nonpublic schools. Socioeconomic and sociological factors that may affect student performance should be considered when interpreting these results.

Average reading scale scores by type of school are presented in figure 2.11. In 2000, the average scores of students attending any of the types of school did not differ significantly from 1998 or from past assessment years.

Comparison of scale score results between the types of schools in 2000 show nonpublic school students outperforming public school students. Students at Catholic and other nonpublic schools also had higher average scores than fourth-graders attending public schools.

Figure 2.11
Scale Score Results by Type of School

Average fourth-grade reading scale scores by type of school: 1992–2000



SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Achievement level results by type of school are presented in figure 2.12. As with scale score results, achievement level results in 2000 show no significant increases or decreases in the percentages of students at or above the *Basic* or *Proficient* achievement levels as compared to past assessment years.

Comparison of 2000 achievement level results between types of schools show fourth-grade students in nonpublic schools outperforming their peers in public schools. All classifications of nonpublic schools had higher percentages of students at or above *Basic*, at *Proficient*, and at or above *Proficient* than did public schools. Among schools classified as "nonpublic" a higher percentage of students attained the *Advanced* level of achievement than did their public school peers. Within the types of nonpublic schools, only those classified as "other nonpublic" had a higher percentage of students at the *Advanced* level than did public schools. The apparent difference between the percentages of students at the *Advanced* level in public and Catholic schools was not statistically significant.

Figure 2.12
Achievement Level Results by Type of School

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by type of school: 1992–2000

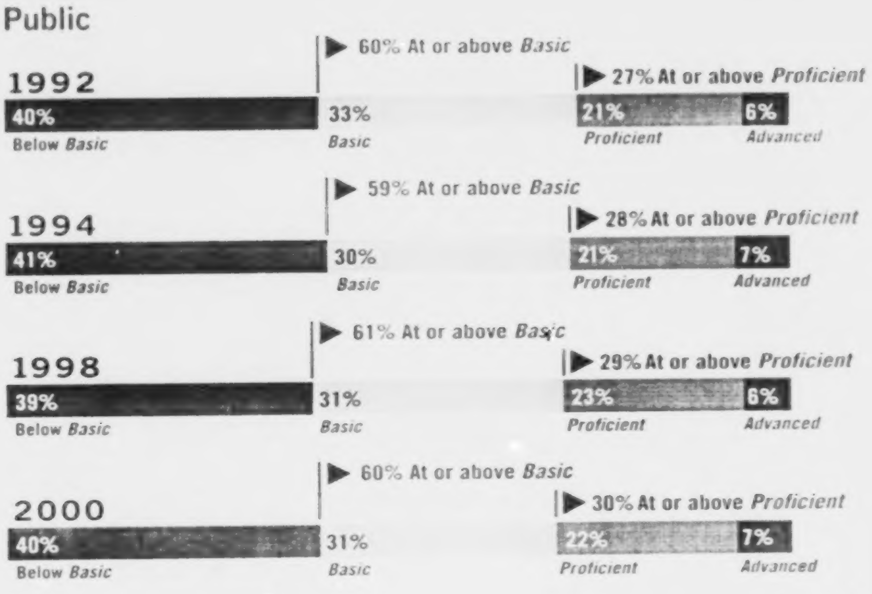


Figure continues on the next page.

Figure 2.12
Achievement Level Results by Type of School (continued)

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by type of school: 1992–2000

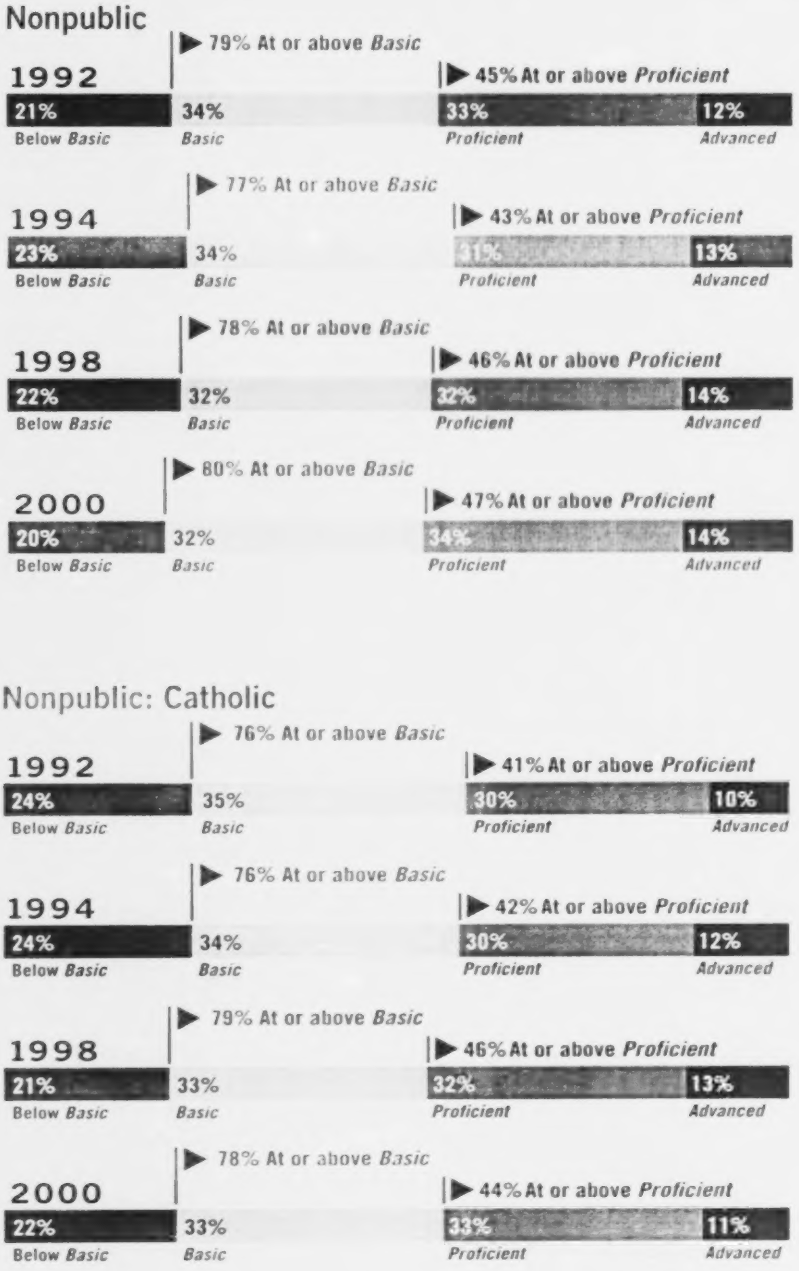
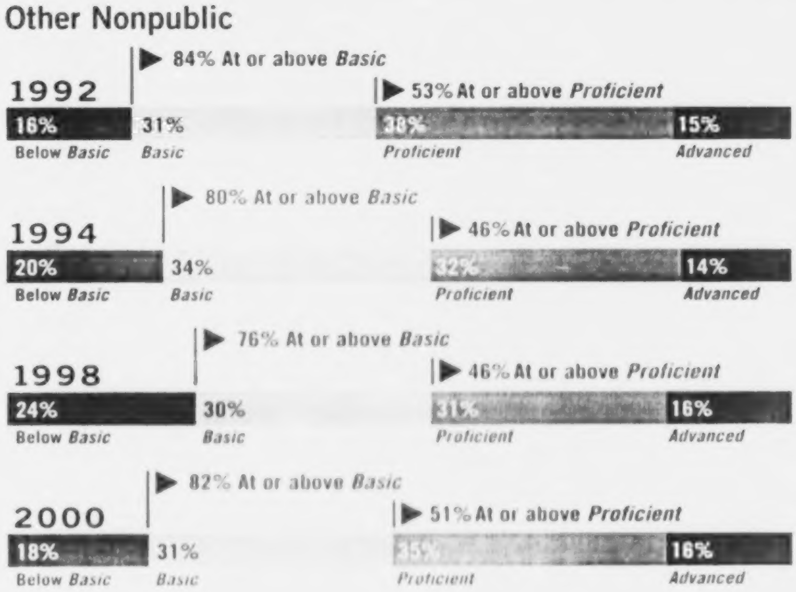


Figure 2.12
Achievement Level Results by Type of School (continued)

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by type of school: 1992–2000



NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Summary

This chapter presented results from the NAEP reading assessment for fourth-grade students in different subgroups. The subgroups examined were gender, race/ethnicity, region of the country, type of location, eligibility for the free/reduced-price lunch program, and type of school.

The following figure displays the major findings presented in this chapter. In each line of the display, the 2000 average reading score of a particular group or the percentage of students at or above the *Proficient* achievement level for a particular group is compared to that in the first assessment year under the current reading framework or to the first year that data was available. Arrows pointing upward (🡅) indicate significant increases, horizontal arrows (➔) indicate no significant change, and arrows pointing downward (🡇) indicate significant decreases. For example, the first section under gender indicates that that there has been no significant increase since 1992 in the average score for either male or female fourth-graders, but there has been an increase since 1992 in the percentage of females at or above the *Proficient* achievement level.

Figure 2.13 Summary of scale scores and achievement levels: 1992–2000

Gender

Female
→ Scale score since 1992
↑ Percentage at or above *Proficient* since 1992

Male
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Race/Ethnicity

White
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Black
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Hispanic
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Asian/Pacific Islander
↑ Scale score since 1992
↑ Percentage at or above *Proficient* since 1992

American Indian
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Region

Northeast
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Southeast
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Central
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

West
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Eligibility for the Free/Reduced-Price Lunch Program

Eligible
→ Scale score since **1998**
→ Percentage at or above *Proficient* since **1998**

Not Eligible
→ Scale score since **1998**
→ Percentage at or above *Proficient* since **1998**

Type of School

Public
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Nonpublic
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Nonpublic: Catholic
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

Other Nonpublic
→ Scale score since 1992
→ Percentage at or above *Proficient* since 1992

NOTE: Years are shown in boldface when the comparison year is different from the initial assessment year because earlier data are not available.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 2000 Reading Assessments.

3

School and Home Contexts for Learning

This chapter presents information on the two contexts that largely form a child's world and contribute most to learning: school and home. Aspects of both school and home are considered because what students do at school and at home contribute, in different ways, to reading development. What students learn and do while at school may be reinforced in the home when literacy-related activities are interwoven into daily life. Both classrooms and homes can be print-rich environments that can accommodate talking about books or reading aloud from them. Quiet time to read on one's own or to write down thoughts about what has been read ideally occurs not only at school, but also at home. Such connections between the worlds of school and home communicate the value of learning and encourage students' reading development.[1]

The information in this chapter is based on students' responses to background questions administered as part of the NAEP 2000 reading assessment. Results from the 2000 assessment are compared to 1992, 1994, and 1998 results. The percentage of students who selected each response option and the average scale scores are presented for each contextual variable reported. Thus, it is possible to examine the relationship between students' school and home experiences and their performance on the NAEP assessment. Readers of this report are reminded that the relationship between a contextual variable and reading performance is not necessarily causal.

Chapter Focus

How do certain school and home factors relate to reading progress?

Have school and home factors related to learning changed since 1992?

Chapter Contents

Students' Reports on Their Reading in and for School

Students' Reports on Literacy with Family and Friends

---

[1] Baker, L. (1999). Opportunities at home and in the community that foster reading engagement. In J. Guthrie and D. Alvermann (Eds.), *Engaged reading: Processes, practices, and policy implications* (pp. 105–133). New York: Teachers College.

Reading in and for School

As with the acquisition of many skills, practice is important to reading development. Fluency, vocabulary, and comprehension may improve in relation to the amount of reading accomplished. While guided oral reading has been shown to effectively enhance a variety of reading skills, that increased silent reading leads to higher reading achievement has not been confirmed by research studies. Yet, it is generally agreed that practice in reading develops better readers.[2]

Pages Read in School and for Homework. Students' reports in 2000 indicate a consistent relationship between the daily amount of reading done in school and for homework and reading performance.

As shown in table 3.1, 80 percent of the students reported reading at least 6 pages daily and 60 percent reported reading 11 or more pages each day. Higher numbers of pages read daily are associated with higher average reading scale scores. Fourth-graders who reported reading 11 or more pages daily had the highest average score, outperforming their peers who reported reading fewer pages.

Looking across assessment years indicates that fourth-graders currently are reading

Table 3.1

Students' reports on the number of pages read each day in school and for homework: 1992–2000

Grade 4

Pages Read in School and for Homework

	1992	1994	1998	2000
11 or more pages				
Percentage of Students	56 *	54 *	57	60
Average Scale Score	222	220	221	222
6 to 10 pages				
Percentage of Students	23 *	23 *	22	20
Average Scale Score	217	214	217	215
5 or fewer pages				
Percentage of Students	21	23 *	21	19
Average Scale Score	203	201	207	202

More students read 11 or more pages each day than in 1992 and 1994.

Students who read most scored highest.

* Significantly different from 2000

NOTE: Percentages may not add up to 100 due to rounding

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments more pages on a daily basis. The percentage of fourth-graders who reported reading the most pages—11 or more daily—was higher in 2000 than the percentages had been in both 1992 and 1994.

---

[2] National Reading Panel. (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research on reading and its implications for reading instruction: Reports of the subgroups* (pp. 3-15–3-38). Washington, DC: National Institute of Child Health & Human Development, National Institutes of Health.

Time Spent Doing Homework. While 2000 results showed students who reported more daily reading in school or for homework outperforming their peers, students' reports on the amount of time spent each day on homework do not necessarily indicate that more time is better. As shown in table 3.2, 72 percent of fourth-grade students reported spending one-half to one hour a day on homework. These students outperformed their peers who reported spending more than one hour on homework each day and those who reported either not doing or not having homework. Students who reported not having any homework or spending more than one hour on homework had similar scores.

Table 3.2

Students' reports on the amount of time spent doing homework each day: 1992–2000

Grade 4

Time Spent Daily Doing Homework

	1992	1994	1998	2000
More than one hour				
Percentage of Students	15	15	16	16
Average Scale Score	208	208	213	212
One hour				
Percentage of Students	28	30	31	29
Average Scale Score	221	218	221	222
One-half hour				
Percentage of Students	39 *	39 *	43	43
Average Scale Score	217	216	219	219
Do not do homework				
Percentage of Students	2	3 *	2	2
Average Scale Score	196	183	188	172
Do not have homework				
Percentage of Students	16 *	13 *	8	10
Average Scale Score	220	216	213	212

The 72% of students who said they spent one-half to one hour on homework each day scored highest.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Some variation can be observed in students' reports about time spent daily on homework since previous NAEP assessments. In 2000, the percentage of fourth-graders who reported spending one-half hour on homework each day was higher in comparison to both 1992 and 1994. While the percentages of students who reported not having homework have been relatively small across assessments, in 2000 this percentage of students was significantly lower than in 1992 and 1994.

Writing About Reading. Writing about reading has become a valued part of instruction that provides learners with ways of thinking about their ideas and exploring content domains.[3] As the framework that underlies the NAEP assessment views reading as an active process, many of the assessment questions ask students to write out their own answers rather than to select a multiple-choice option. How prepared are fourth-graders to write out their own answer about reading by having been asked to do so in school?

Fourth-graders who took the NAEP reading assessment were asked how frequently during the school year they had been asked to write long answers to questions on tests and assignments that involved reading. As shown in table 3.3, 81 percent of fourth-grade students reported writing long answers about reading at least once or twice a month. Students who reported being asked to produce such answers on a monthly basis had a higher average score than fourth-graders asked to do so weekly and those asked to do so less frequently. While 2000 results show highest performance among students who are asked to engage in this activity once or twice monthly, students asked to write out long answers weekly outperformed their peers who reported doing so yearly or never or hardly ever. Lowest reading performance was associated with never or hardly ever writing long answers about reading.

Increase in the frequency of this activity at grade 4 observed in 1998 was maintained by the most recent student reports. In 2000, the percentage of students who reported being asked to write long answers to questions that involved reading at least weekly was again higher in comparison to 1994. A decrease since 1994 was observed in the small percentage of students reporting never being asked to write long answers to questions about reading.

---

[3] McGinley, W. & Tierney, R. J. (1989). Traversing the topical landscape. *Written Communications* 6 (3), 243–269.

Table 3.3

Students' reports on how often they write long answers to questions on tests or assignments that involved reading: 1992–2000

Grade 4

Writing About Reading

	1992	1994	1998	2000
At least once a week				
Percentage of Students	51	48 *	53	53
Average Scale Score	220	217	218	217
Once or twice a month				
Percentage of Students	28	31 *	30	28
Average Scale Score	221	221	223	225
Once or twice a year				
Percentage of Students	13	12	10	11
Average Scale Score	209	209	212	210
Never or hardly ever				
Percentage of Students	9	9 *	8	8
Average Scale Score	202	198	199	199

The 81% of students who said they write about reading at least monthly scored highest.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Teachers' Help with Words. The last school factors examined in this chapter look at the frequency of two distinctly different types of word-level activity and their relationship to fourth-grade reading performance. In 1998 and 2000, students who took the NAEP assessment were asked to report how often their teacher helped them to break words into parts and how often their teacher helped them to understand new words. Results indicate that these two activities had a different relationship to performance.

Help with breaking words into parts had a consistent negative relation to fourth-grade reading performance as demonstrated on the NAEP assessment. As shown in table 3.4, the more frequently students received this help the lower their average score; whereas fourth graders who reported that their teachers never helped them break words into parts had the highest average score. These results are consistent with research findings that have shown this type of word-level reading instruction most effective in kindergarten and the 1st grade before children have learned to read independently and also effective in helping older struggling readers.[4] The 53 percent of fourth-graders who reported never or rarely needing this type of help demonstrated the highest comprehension skills, outperforming those fourth-graders who were still focusing more on recognizing individual words rather than on comprehending the meaning of the text as a whole. If by the fourth grade students still require daily or weekly help breaking words into parts, it may likely be that these students are having difficulty learning to read.

While breaking words into parts is useful in the early stages of learning to read, learning new words is a part of reading at all stages of development. By fourth grade, students are beginning to explore content areas and may encounter vocabulary that is specific to that content. Vocabulary instruction, either by explaining new words prior to reading or by explaining them in the context of the reading experience, can facilitate fourth-graders' comprehension of new content material.[5] Broadly stated, help breaking words into parts is associated with learning to read and help with new words is more closely linked with reading to learn. As shown in table 3.4, students who reported receiving a moderate amount of help with new words—weekly or monthly—had higher average scores than fourth-grade students who reported receiving this help either every day or never/hardly ever. It may be that the fourth-graders who required help weekly or monthly are reading independently and encountering texts that challenge them with some new words.

---

[4] National Reading Panel. (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (pp. 2-81–2-130). Washington, DC: National Institute of Child Health & Human Development, National Institutes of Health.

Foorman, B., Francis, D., Fletcher, J., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90,* 37–55.

[5] National Reading Panel (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (pp. 4-15–4-25). Washington, DC: National Institute of Child Health & Human Development, National Institutes of Health.

Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *Elementary School Journal, 96* (4), 415-422.

Table 3.4

Students' reports on how often their teachers help them break words into parts and help them understand new words: 1998–2000

Teachers' Help with Words

Grade 4

	1998	2000
How often their teachers help them break words into parts		
Every day		
Percentage of Students	25	25
Average Scale Score	210	209
Once or twice a week		
Percentage of Students	23	22
Average Scale Score	217	217
Never or hardly ever		
Percentage of Students	52	53
Average Scale Score	226	226

Students who said they never or rarely received help breaking words into parts scored highest.

	1998	2000
How often their teachers help them understand new words		
Every day		
Percentage of Students	49	48
Average Scale Score	217	217
Once or twice a week		
Percentage of Students	24	23
Average Scale Score	224	224
Once or twice a month		
Percentage of Students	14	14
Average Scale Score	223	224
Never or hardly ever		
Percentage of Students	12 *	14
Average Scale Score	219	216

Students who said their teacher helped them with new words weekly or monthly scored highest.

* Significantly different from 2000

NOTE: Percentages may not add up to 100 due to rounding

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

BEST COPY AVAILABLE

Literacy with Family and Friends

Learning is not limited to the classroom but is a process that continues in, and is shaped by, home environments and social interactions. Families use print for various activities on a daily basis, and different cultural groups have unique ways of integrating oral and written language with daily social life.[6] Occasions for learning and for enhancing reading skills ideally coincide with the behaviors of everyday life. This section of the chapter looks at literacy-related activities that occur outside of school and their relationship to student performance on the 2000 NAEP reading assessment.

Reading for Fun. As noted earlier, the amount of reading students do in and for school on a daily basis had a positive relationship to performance on the NAEP reading assessment. If students, however, regard reading only as a school-related activity, as a duty rather than a pleasure, their future prospects for reading to understand themselves and the world are limited. For reading on one's own not only extends comprehension skills, but also enhances the understanding of what happens in life. While reading stories, for instance, children use literature to make connections and comparisons, providing them with a perspective that may exceed the boundaries of their immediate experience.[7] When the act of reading extends beyond the schoolroom and becomes part of daily life, ongoing literacy is on its way to becoming a reality.

As shown in table 3.5, 75 percent of fourth-grade students reported reading for fun at least once or twice a week. More frequent reading for fun showed a positive relationship to scores. Students who reported reading for fun daily had higher average scores than students who reported reading for fun less frequently.

While students' reports indicate relatively stable percentages of students engaged in reading on their own across assessment years, the percentage of students in 2000 who reported never reading for fun on their own was significantly higher than in 1994. This contrasts with students' reports on school-related reading, noted earlier in this chapter, which showed higher percentages of students reading more pages daily in school and for homework than in 1992 and 1994.

---

[6] Gee, J. (1990). *Social linguistics and literacies.* London: Falmer.

[7] Wolf, S.A. & Heath, S.B. (1992). *The braid of literature: Children's world of reading.* Cambridge, MA: Harvard University Press.

Table 3.5

Students' reports on how often they read for fun on their own time: 1992–2000

Grade 4

	1992	1994	1998	2000
Every day				
Percentage of Students	44	45	43	43
Average Scale Score	223	223	222	223
Once or twice a week				
Percentage of Students	32	32	32	32
Average Scale Score	218	213	219	218
Once or twice a month				
Percentage of Students	12	12	12	12
Average Scale Score	210	208	216	216
Never or hardly ever				
Percentage of Students	13	12 *	13	14
Average Scale Score	199	197	203	202

Reading for Fun

Students who read for fun every day scored highest.

Three quarters of the students said they read for fun at least weekly.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Discussing Studies and Talking About Reading. Reading, whether for school or for fun, is only one aspect of literacy development. Equally important is the social dimension that brings books off the shelves and out of bookbags into the dynamics of daily life. Social interactions that provide students with the opportunity to talk with others about their reading enhance literacy skills and encourage further reading activity.[8] Fourth-grade students in the NAEP assessment were asked to indicate how frequently they discussed their studies at home or talked about reading with family and friends. As shown in table 3.6, students' reports suggest a distinction between the two types of discourse. Students who discussed their studies at home, however frequently, had higher average reading scores than students who reported never discussing their studies at home. There was no significant difference between the scores of those who did so daily, weekly, or monthly.

Somewhat more frequent talk with family and friends about their reading showed a positive relation to student performance. Fourth-graders who responded that they engaged in conversation about their reading with family or friends on a weekly basis had a higher average score than students who reported engaging in such talk daily, monthly, or never. Students who reported never or rarely doing so had the lowest average score.

The percentages of students reporting various frequencies of talking about studies or their reading outside of school have remained relatively stable across assessment years. No significant change was observed in student reports of how often they discuss studies at home. An increase since 1994 was observed, however, in the percentage of students who reported never or hardly ever talking about their reading with family or friends. Considering the positive relationship this activity has with reading achievement, it is disappointing to note that nearly one-quarter of students reported never or hardly ever doing so.

---

[8] Guthrie, J., Shafer, W.D., Yang, Y.Y., & Afflerbach, P. (1995). Relationships of instruction on reading: An exploration of social, cognitive and instructional connections. *Reading Research Quarterly, 30*, 8–25.

Table 3.6

Students' reports on how often they discuss their studies at home and talk about reading with their family and friends: 1992–2000

Grade 4

Discussing Studies and Talking About Reading

	1992	1994	1998	2000
Discuss studies at home				
Almost every day				
Percentage of Students	54	55	54	54
Average Scale Score	221	219	220	221
Once or twice a week				
Percentage of Students	22	22	23	23
Average Scale Score	220	215	222	219
Once or twice a month				
Percentage of Students	6	6	6	6
Average Scale Score	215	208	213	217
Never or hardly ever				
Percentage of Students	17	17	18	17
Average Scale Score	202	199	205	201

The 83% of students who said they discussed their studies at home however frequently outperformed students who never or rarely do so.

	1992	1994	1998	2000
Talk about reading with family or friends				
Almost every day				
Percentage of Students	26	28	27	27
Average Scale Score	215	213	211	213
Once or twice a week				
Percentage of Students	36	36	35	34
Average Scale Score	224	223	223	227
Once or twice a month				
Percentage of Students	15	15	15	15
Average Scale Score	219	214	222	220
Never or hardly ever				
Percentage of Students	23	21 *	23	24
Average Scale Score	209	207	214	209

75% of students said they talked about reading with family and friends at least monthly. Students who said they did so weekly scored highest.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

BEST COPY AVAILABLE

Reading Materials in the Home. One indication that a household values literacy as an inherent part of everyday life is the presence of a variety of reading materials. Students who participated in the NAEP reading assessment were asked to indicate if their family regularly received a newspaper or magazines, if there was an encyclopedia at home, and to approximate how many books were in the home. Students were classified as having books in the home if they reported having more than 25 books. Reports on these four different types of reading materials are presented in table 3.7.

Results of the 2000 NAEP reading assessment are consistent with the findings of previous assessments that showed higher average scores among students who reported having more types of reading materials at home.[9] Students who reported having all four types of materials in their home had higher average scores than students who reported having three types of materials; those who indicated the presence of three types had in turn a higher average score than students indicating the presence of two or fewer types of reading materials.

Perhaps the availability of newspapers, magazines, and encyclopedias in electronic form has influenced the number of different types of reading materials in the home reported by students. Although there was no significant change since 1998, a lower percentage of students in 2000 reported having all four types of reading materials at home in comparison to 1994 reports.

---

[9] Donahue, P.D., Voekl, K.R., Campbell, J.R. & Mazzeo, J. (1999). *NAEP's 1998 reading report card for the nation and the states.* National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Campbell, J.R., Donahue, P.D., Reese, C.M. & Phillips, G.W. (1996). *NAEP's 1994 reading report card for the nation and the states.* National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Table 3.7

Students' reports on the number of different types of reading materials in the home: 1992–2000

Grade 4

Reading Materials in the Home

	1992	1994	1998	2000
Four				
Percentage of Students	37	38 *	37	34
Average Scale Score	226	227	228	229
Three				
Percentage of Students	32	34	33	34
Average Scale Score	219	216	220	219
Two or fewer				
Percentage of Students	31	29 *	30	32
Average Scale Score	204	197	204	203

Two-thirds of students said they had three or four types of reading materials in their homes. Students who had all four types scored highest.

* Significantly different from 2000

NOTE: Percentages may not add up to 100 due to rounding

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments

Time Spent Watching Television. The last home variable to be considered in this chapter is the amount of time students reported watching television daily. While the availability of technologies including television and computers have been associated with reading engagement,[10] and although many educationally oriented programs are aired, concern remains that television watching interferes with the time students spend on more active literacy pursuits.

Students' reports on the amount of time they spend watching television are presented in table 3.8. Results from the NAEP 2000 reading assessment once again show that watching many hours of television daily has a negative relationship to reading performance. Students who reported watching the most television, six hours or more, had the lowest average reading score, and those who reported watching four to five had the next lowest score. Fourth-grade students who reported watching less television, either two or three hours or an hour or less daily, had higher and similar scores.

A clear trend in whether students are watching more or less television is not evident. Although a significantly higher percentage of fourth-grade students in 2000 than in 1998 reported watching six or more hours of television daily, this percentage is lower than it was in 1994. However, results also show increases in the percentages of students watching three or fewer hours of television daily. The percentage of students in 2000 who reported watching two or three hours was higher than in 1994, and the percentage who reported watching one hour or less was higher in comparison to both 1992 and 1994.

---

[10] Baker, L. (1999). Opportunities at home and in the community that foster reading engagement. In J. Guthrie & D. Alvermann (Eds.) *Engaged reading: Processes, practices and policy implications.* (pp. 105–33), New York: Teachers College.

Table 3.8

Students' reports on the amount of time spent watching television each day: 1992–2000

Grade 4

Time Spent Daily Watching Television

	1992	1994	1998	2000
Six hours or more				
Percentage of Students	20	21 *	16 *	18
Average Scale Score	199	194	198	196
Four or five hours				
Percentage of Students	22 *	22 *	19 *	17
Average Scale Score	216	216	216	213
Two or three hours				
Percentage of Students	40	38 *	41	40
Average Scale Score	224	222	223	224
One hour or less				
Percentage of Students	19 *	19 *	24	25
Average Scale Score	221	220	222	224

Since 1992, fewer students said they are watching four or five hours of television each day.

Students who said they watch less television each day scored highest.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

BEST COPY AVAILABLE

Summary

While no causal relationship can be made between student reports on school and home experiences and their reading performance, results presented in this chapter do suggest that active engagement with words and meanings may be an essential factor for literacy development.

Students who reported reading more pages daily in school and for homework and students who reported being asked to write about their reading were higher achievers. Activities outside of school, such as reading for fun, discussing studies at home, or talking about reading with family and friends also showed a positive influence on fourth-grade reading achievement as measured by the NAEP assessment.

Considering that students' reports indicating more involvement with language at school and at home were positively related to scale scores, it is not surprising that having access to more types of literacy materials in the home—materials that might stimulate reading activity—were also reported by higher-achieving students. The inverse of active engagement in literacy-related activities is reflected in students' reports on television viewing. Students who reported watching the most television daily were the lowest achievers.

Some of the factors that had a positive relationship with students' scores on the NAEP reading assessment have shown encouraging change across assessment years. In comparison to 1992 and 1994, higher percentages of students in 2000 reported reading more pages in school and for homework on a daily basis; also, in comparison to 1994, higher percentages of students reported being asked to write about their reading weekly or monthly.

While these school-related factors indicate students' increased involvement with language activity, some home factors related to literacy development do not reflect a positive trend. In comparison to 1994, small increases can be observed in the percentages of students who reported never or hardly ever reading for fun on their own time or talking about their reading with family and friends. Although two-thirds of students reported having three or four types of reading materials in their homes, more students than in 1994 reported having two or fewer types. No clear pattern over time was observed in students reports on how much television they watch daily; however, it is encouraging that in comparison to 1992 and 1994 a higher percentage of fourth-graders reported watching one hour or less of television each day.

Overall, it may be said from the results in this chapter that high percentages of fourth-grade students are engaged in literacy-related activities on a fairly regular basis. For example, three-quarters of fourth-graders reported reading for fun weekly and talking about reading with family and friends at least monthly; and over three-quarters of the students reported discussing their studies at home. Students' reports on school activities show 60 percent of fourth-graders reading eleven or more pages daily in school or for homework and 72 percent spending one-half to one hour daily on homework.

4 Becoming a More Inclusive National Assessment

Chapter Focus

How would the NAEP results differ if accommodations were permitted for special needs students?

Chapter Contents

Two sets of 2000 NAEP Reading Results

Results for the Nation

Results by Gender

Results by Race/Ethnicity

Legislation at the federal level now mandates the inclusion of all students in large-scale academic assessments.[1] As a consequence, the majority of states have assessment programs that must make provisions to include special-needs students—provisions that include the allowance of testing accommodations when appropriate. Assessing as representative a sample of the nation's students as possible is particularly important for NAEP's mission to serve as a key indicator of the academic achievement of the nation's students. This mission can be satisfactorily accomplished only if the assessment results include data gathered from all groups of students, including those classified as having special needs.

Although the intent of NAEP has consistently been to include special-needs students in its assessments to the fullest degree possible, the implementation of the assessment has always resulted in some exclusion of students with disabilities and students with limited English proficiency who could not be assessed meaningfully without accommodations. Participating schools have been permitted to exclude certain students who have been classified as having a disability under the Individuals with Disabilities Education Act, based upon their Individualized Education Programs (IEP) and Section 504 of the Rehabilitation Act of 1973. Similarly, schools have been permitted to exclude some students they identify as being limited English proficient. Exclusion decisions are made in accordance with explicit criteria provided by the NAEP program.

---

[1] Goals 2000, Elementary and Secondary Education Act (ESEA), Improving America's Schools Act (IASA), Individuals with Disabilities Education Act (IDEA). See also: Title VI of the Civil Rights Act, Equal Educational Opportunities Act, Section 504 of the Rehabilitation Act.

In order to move the NAEP assessments toward more inclusive samples, the NAEP program began to explore the use of accommodations with special-needs students during the 1996 and 1998 assessments. An additional impetus for this change was to keep NAEP consistent with state and district testing policies that increasingly offered accommodations so that more special-needs students could be assessed. In both 1996 and 1998, the national NAEP sample was split so that part of the schools sampled were permitted to provide accommodations to special-needs students and the other part was not. This sample design made it possible to study the effects on NAEP results of including special-needs students in the assessments under alternate testing conditions. Technical research papers have been published with the results of these comparisons.[2] Based on the outcomes of these technical analyses, the 1998 results of those NAEP assessments that used new test frameworks (writing and civics), and hence also began new trend lines, were reported with the inclusion of data from accommodated special-needs students.

The results presented in the 1998 reading report card included the performance of students with disabilities (SD) and those with limited English proficiency (LEP) who were assessed without the possibility of accommodations. The results did not include the performance of students for whom accommodations were permitted because of the need to preserve comparability with the results from 1992 and 1994. Students in those earlier assessments had not had accommodations available to them. However, in both the 1998 and 2000 reading assessments, the NAEP program used the split-sample design, so that trends in students' reading achievement could be reported across all the assessment years and, at the same time, the program could continue to examine the effects of including students tested with accommodations.

Two Sets of 2000 NAEP Reading Results

This report card is the first to display two different sets of NAEP results based on the split-sample design: (1) those that reflect the performance of regular and special-needs students when accommodations were not permitted, and (2) those that reflect the performance of regular and special-needs students—both those who were accommodated and those who could test without accommodations—when accommodations were permitted. It should be noted that accommodated students make up a small proportion (about 3 percent) of the total weighted number of students assessed (see table A.6 in appendix A for details). Making accommodations available may change the overall assessment results in subtle and different ways. For example, when accommodations are permitted, there may be some occurrences of students being accommodated who may have taken the test under standard conditions if accommodations were not permitted. This could lead to an overall increase in the average assessment results, if it can be assumed that accommodations increase special-needs students' performance. Conversely, when accommodations are permitted, special-needs students who could not have been tested without accommodations could be included in the sample. Assuming that these are generally lower-performing students, their inclusion in the sample—even with accommodations—may result in an overall lower average score. The findings in the NAEP research reports cited previously suggest that these two opposite influences on the total assessment results may balance out to result in little change in the overall assessment averages.

---

[2] Olson, J.F. and Goldstein, A. A. (1997). *The inclusion of students with disabilities and limited English proficient students in large-scale assessments: A summary of recent progress.* (NCES Publication No. 97–482). Washington, DC: National Center for Education Statistics.

Mazzeo, J., Carlson, J.E., Voelkl, K.E., & Lutkus, A. D. (1999). *Increasing the participation of special needs students in NAEP: A report on 1996 research activities.* (NCES Publication No. 2000–473). Washington, DC: National Center for Education Statistics.

The first three chapters of this report are based on the first set of results (no accommodations offered). This final chapter presents an overview of the second set of results—results that include students who were provided accommodations during the test administration. By including these results, the NAEP program begins a phased transition toward a more inclusive reporting sample. Future assessment results will be based solely on a student and school sample in which accommodations are permitted.

The two sets of results presented in this chapter were obtained by administering the assessment to a nationally representative sample of students and schools. In half of the schools sampled, no accommodations were permitted; all students were tested under the same conditions that were the basis for reporting results from the 1992, 1994, and 1998 NAEP reading assessments. In the other half of the schools sampled, accommodations were permitted for students with disabilities and limited English proficient students who normally receive accommodations in their district or state testing programs. Most accommodations that schools routinely provide for their own testing programs were permitted. The permitted accommodations included, but were not limited to:

- one-on-one testing,
- small-group testing,
- extended time,
- oral reading of directions,
- signing of directions,
- use of magnifying equipment, and,
- use of an aide for transcribing responses.

The program did not allow some of the accommodations that are permitted in certain states. In particular, some states allow questions and, in some instances, reading passages to be read aloud to the students. Such "read-aloud accommodations" were viewed by the program as changing the nature of the construct being measured and, hence, were not permitted. Because the NAEP program considers the domain of its reading assessment as "reading in English," no attempt was made to provide an alternate-language version of the instrument and the use of bilingual dictionaries was not allowed. (See appendix A, table A.7 for greater detail on numbers and percentages of students accommodated by accommodation type in the 1998 and 2000 assessments.) The "read-aloud" accommodation is permitted, however, in NAEP subjects other than reading.

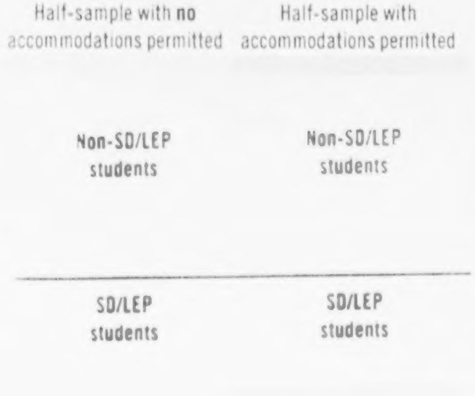
Figure 4.1 provides a visual representation of how the two sets of results were based on the two half-samples in 1998 and 2000. Included in both sets of results (accommodations not permitted and accommodations permitted) are those students from both half-samples of schools who were not identified as having a disability (SD) or limited English proficiency (LEP). In addition, the first set of results (accommodations not permitted) includes SD and LEP students from the half-sample of schools where accommodations were not permitted (see middle portion of Figure 4.1). This is the set of results that allows for trend comparisons back to 1992 and are reported in the first three chapters of this report.

The second set of results, accommodations permitted (see bottom portion of Figure 4.1), includes SD and LEP students from the half-sample of schools where accommodations were permitted. This is the set of results that form the new, more inclusive baseline for future reporting of trend comparisons for the NAEP reading assessment.

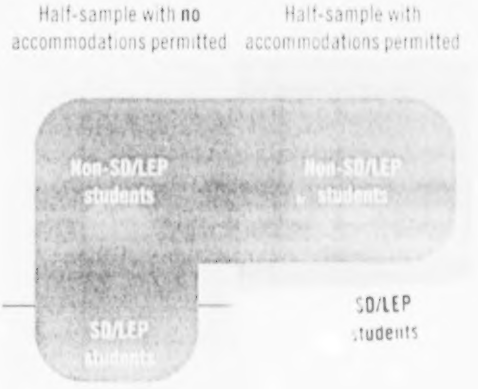
In the NAEP 2000 samples where accommodations were not permitted, 15 percent of the students were identified by their schools as having special needs (i.e., either as students with disabilities or limited English proficient students). In the other half-sample where accommodations were offered, 17 percent of the students were identified as having special needs. In the sample where accommodations were not permitted, 48 percent of the special-needs students (7 percent of all students, see appendix A, table A.5) were excluded from NAEP testing by their schools. In the sample where accommodations were offered, only 35 percent of the special-needs students were excluded from testing (6 percent of the total sample). Thus, as seen with other NAEP subjects, offering accommodations leads to greater inclusion of special-needs students.

Figure 4.1

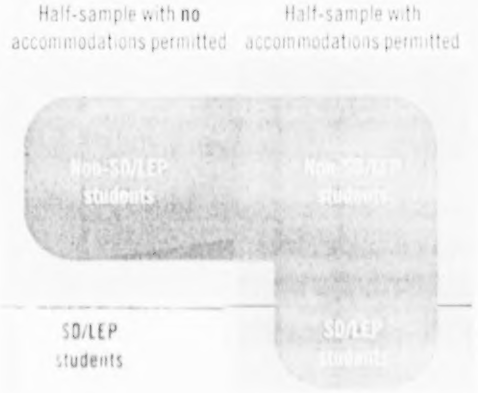
The two sets of NAEP results based on a split-sample design


Half-sample with no accommodations permitted
Half-sample with accommodations permitted
Non-SD/LEP students
Non-SD/LEP students
SD/LEP students
SD/LEP students

Split-sample design
The national sample was split. In half of the schools, accommodations were not permitted for students with disabilities (SD) and students with limited English proficiency (LEP). In the other half of the schools, accommodations were permitted for SD and LEP students who routinely received them in their school assessments.


Half-sample with no accommodations permitted
Half-sample with accommodations permitted
Non-SD/LEP students
Non-SD/LEP students
SD/LEP students
SD/LEP students

Accommodations-not-permitted results
The accommodations-not-permitted results include the performance of students from both half-samples who were not classified as SD or LEP and the performance of SD and LEP students from the half-sample in which no accommodations were permitted.


Half-sample with no accommodations permitted
Half-sample with accommodations permitted
Non-SD/LEP students
Non-SD/LEP students
SD/LEP students
SD/LEP students

Accommodations-permitted results
The accommodations-permitted results also include the performance of students from both half-samples who were not classified as SD or LEP; however, the SD and LEP students whose performance is included in this set of results were from the half-sample in which accommodations were permitted. Since students who required testing accommodations could be assessed and represented in the overall results, it was anticipated that these results would include more special-needs students and reflect a more inclusive sample.

The focus of this chapter is a comparison of data from the two sets of results: (1) when accommodations were not permitted, and (2) when accommodations were permitted. Because the split-sample design was used in both 1998 and 2000 for the NAEP reading assessment, both sets of results are presented for both years. Overall results are provided for the nation, as well as for student subgroups by gender and by race/ethnicity. These results are discussed in terms of statistically significant differences between the two sets of results in each year, changes between assessment years, and differences between subgroups of students within each set of results. Throughout this chapter, the assessment results that include SD and LEP students for whom accommodations were not permitted will be referred to as the "accommodations-not-permitted" average score. The set of results that includes SD and LEP students for whom accommodations were permitted will be referred to as the "accommodations-permitted" average score.

Results for the Nation:

Accommodations Not Permitted and Accommodations Permitted

Table 4.1 displays the average reading scale scores for the nation in 1998 and 2000 for the two sets of results: (1) accommodations not permitted, and (2) accommodations permitted. In the 1998 reading assessment at grade 4, there was no statistically significant difference between the two average scores. However, in 2000 the accommodations-permitted average score was two points lower than the accommodations-not-permitted average score. An important comparative context for this finding is that NAEP's previous research in the science and mathematics subjects in the 1996 assessments found no significant differences in the average scores of the two sets of results at grades 4, 8, or 12.[3]

Table 4.1

Average score by type of results: 1998 and 2000

	Accommodations not permitted	Accommodations permitted
1998	217	216
2000	217	215*

*Significantly different from the sample where accommodations were not permitted.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

---

[3] Mazzeo, J., Carlson, J.E., Voelkl, K.E., & Lutkus, A. D. (1999). *Increasing the participation of special needs students in NAEP: A report on 1996 research activities.* (NCES Publication No. 2000–473). Washington, DC: National Center for Education Statistics.

As noted in the introduction to this chapter, NAEP has always sought to include special-needs students proportional to their representation in the U.S. population. Offering accommodations tends to reduce exclusion rates for special-needs students and therefore allows NAEP to offer a fairer and more accurate picture of the status of American education. Because special-needs students are typically "classified" as eligible for special educational services after having shown some difficulty in the regular learning environment, it may be assumed that the academic achievement of special-needs students will be lower than that of students without such needs. This assumption appears to have been justified in the observed difference between the two sets of fourth-grade reading results in 2000, where the accommodations-permitted results, which included slightly more special-needs students because of the availability of accommodations, were lower than the accommodations-not-permitted results. It is important to examine the percentages of students attaining the NAEP achievement levels to see if the percentages in the lower performance categories (i.e., below *Basic* and *Basic*) were higher in the set of results where accommodations were offered.

Table 4.2 shows the percentages of students attaining each of the achievement levels. The percentages are similar across the two sets of 2000 results; the differences between the accommodations-not-permitted and the accommodations-permitted results were not significantly different. Similarly, the achievement level distributions for the two sets of results in the 1998 reading assessment showed no significant differences.

Table 4.2

Percentages of students attaining reading achievement levels by type of results (Accommodations Not Permitted and Accommodations Permitted): 1998 and 2000

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
1998							
	Not permitted	38	32	24	7	62	31
	Permitted	39	31	23	8	61	31
2000							
	Not permitted	37	31	24	8	63	32
	Permitted	39	30	23	7	61	31

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Results by Gender:

Accommodations Not Permitted and Accommodations Permitted

In past NAEP assessments of reading, female students outperformed male students.[4] The average reading scale scores by gender for both sets of results in 1998 and 2000 are provided in table C.30 in appendix C. The small differences within gender groups between the accommodations-not-permitted and accommodations-permitted average scores were not significant for either 1998 or 2000.

Female students outperformed male students by 10 points in 2000, regardless of whether accommodations were permitted. In addition, the gap in average scores between males and females was larger in 2000 than in 1998 in both sets of results.

For male students, the accommodations-permitted average score declined from 214 to 210 between 1998 and 2000. However, the small decline between 1998 and 2000 in the accommodations-not-permitted average score was not statistically significant. For female students, no significant score difference between 1998 and 2000 is evident in either set of results.

The percentages of male and female students attaining the *Basic, Proficient,* and *Advanced* levels are provided in table C.31 in appendix C. Comparing the two sets of results both in 1998 and 2000, no significant differences were found in the percentages of students attaining each of the achievement levels. Neither set of results shows significant change between 1998 and 2000 for either males or females.

Results by Race/Ethnicity:

Accommodations Not Permitted and Accommodations Permitted

NAEP assessments across academic subjects have typically reported large score differences according to race and ethnic group membership. To the extent that students with disabilities or limited English proficient students may be overrepresented in a particular racial or ethnic group, that group's assessment scores may decrease. Table C.32 in appendix C provides the average reading scale scores for each of the race/ethnicity categories for the two sets of results in 1998 and 2000. Of the race/ethnicity categories, "Hispanic" was the only one in which a significant difference was observed between the accommodations-not-permitted and accommodations-permitted average scores. In the 2000 assessment, the results for Hispanic students that included those who received accommodations had an average scale score that was seven points lower than the results for Hispanic students that did not include those who received accommodations. In 1998, however, the difference between these two sets of results for Hispanic students was smaller and was not statistically significant.

---

[4] Mullis, I.V.S., Campbell, J. R., & Farstrup, A. E., (1993). *NAEP 1992 reading report card for the nation and the states.* Washington, DC: National Center for Education Statistics.

Campbell, J. R., Donahue, P. D., Reese, C. M., & Phillips, G.W. (1996) *NAEP 1994 reading report card for the nation and the states.* Washington, DC: National Center for Education Statistics.

Donahue, P.L., Voelkl, K.R., Campbell, J.R., & Mazzeo, J. (1999). *The NAEP 1998 reading report card for the nation and the states.* Washington, DC: National Center for Education Statistics.

As noted in chapter 2, a pattern of performance differences by race/ethnicity can be seen in the accommodations-not-permitted results in 2000. Both white and Asian/Pacific Islander students scored higher than black, Hispanic, or American Indian students. The same pattern can be observed in the accommodations-permitted results. The only difference noted in the performance by ethnicity pattern between the two sets of results was that in the accommodations-permitted results, American Indian students scored higher than Hispanic students. This was not the case in the accommodations-not-permitted results. The percentages of students in each race/ethnicity category who attained the *Basic, Proficient,* and *Advanced* levels are provided in table C.33 in appendix C. No significant differences were found between the accommodations-not-permitted results and the accommodations-permitted results for the percentages of students attaining each of the achievement levels in 1998 or 2000. While, as noted above, there was a scale score difference for Hispanic students between the two sets of results in the 2000 assessment, a similar statistically significant difference was not observed for Hispanic students in terms of the achievement levels.

Summary

This chapter compared the reading performance of the nation's fourth-graders when accommodations were not permitted and when accommodations were permitted for special-needs students in the 1998 and 2000 NAEP assessments. Both samples being compared in this chapter included non-special-needs students and special-needs students who were tested without accommodations. In the sample with accommodations permitted, those students who customarily receive accommodations in their school's testing received them in the NAEP testing. While NAEP's previous research in the 1996 science and mathematics assessments found no significant differences between the two types of results at grades 4, 8, or 12, in the 2000 reading assessment the accommodations-permitted average score at grade 4 was lower than the accommodations-not-permitted average score.[5]

Within the accommodations-permitted results, the average score for male students declined between 1998 and 2000. For Hispanic students, the accommodations-permitted average score was lower than the accommodations-not-permitted average score in 2000. The comparable difference between the two sets of results for Hispanic students in 1998 was not statistically significant. No significant difference between the two sets of results was evident for any other racial/ethnic subgroup of students in 1998 or 2000.

---

[5] Mazzeo, J., Carlson, J.E., Voelkl, K.E., & Lutkus, A.D. (1999). *Increasing the participation of special needs students in NAEP: A report on 1996 research activities.* (NCES Publication No. 2000–473). Washington, DC: National Center for Education Statistics.

Appendix A

Overview of Procedures Used for the NAEP 2000 Grade 4 Reading Assessment

This appendix provides an overview of the NAEP 2000 grade 4 reading assessment's primary components—framework, development, administration, scoring, and analysis. A more extensive review of the procedures and methods used in the reading assessment will be included in the forthcoming *NAEP 2000 Technical Report*.

Appendix Contents

The Assessment
The Sample
Data Collection
Data Analysis
NAEP Reporting Groups
Cautions in Interpretations

Appendix Focus

Technical Aspects of the NAEP 2000 Reading Assessment.

The NAEP 2000 Grade 4 Reading Assessment

The reading framework underlying the NAEP 2000 assessment originated from a consensus among an array of individuals interested in education about the nature of reading comprehension. This framework was also used in the 1992, 1994, and 1998 reading assessments, permitting analyses of trends in reading performance.

The framework's purpose was to provide a definition of reading on which to base the NAEP assessment. Developing this framework and the specifications that guided development of the assessment involved the critical input of many people including representatives of national education organizations, teachers, parents, policymakers, business leaders, and members of the general public. This consensus process was managed by the Council of Chief State School Officers for the National Assessment Governing Board.

The framework sets forth a broad definition of "reading literacy" that entails not only being able to read, but also knowing when to read, how to read, and how to reflect on what has been read. In addition, the framework views reading as an interactive process in which the reader's abilities, interests, and prior knowledge interact with the text and the context of the reading situation as meaning construction occurs.

The aspects of grade 4 reading literacy described by the reading framework, including purposes for reading and reading stances, are presented in figure A.1. This figure also provides examples of the types of questions that were used to assess the different purposes for reading via the four reading stances.

Figure A.1

1992, 1994, 1998, and 2000 NAEP framework—aspects of grade 4 reading literacy

	Constructing, extending, and examining meaning			
	Initial understanding *Requires the reader to provide an initial impression or unreflected understanding of what was read.*	**Developing an interpretation** *Requires the reader to go beyond the initial impression to develop a more complete understanding of what was read.*	**Personal reflection and response** *Requires the reader to connect knowledge from the text with his/her own personal background knowledge. The focus here is on how the text relates to personal knowledge.*	**Demonstrating a critical stance** *Requires the reader to stand apart from the text and consider it.*
Reading for literary experience	What is the story/plot about?	How did the plot develop?	How did this character change your idea of ________?	Rewrite this story with ________ as a setting or ________ as a character.
	How would you describe the main character?	How did this character change from the beginning to the end of the story?	Is this story similar to or different from your own experience?	How does this author's use of ________ (irony, personification, humor) contribute to ________?
Reading to gain information	What does this article tell you about ________?	What caused this event?	What current event does this remind you of?	How useful would this article be for ________? Explain.
	What does the author think about this topic?	In what ways are these ideas important to the topic or theme?	Does this description fit what you know about ________? Why?	What could be added to improve the author's argument?

SOURCE: National Assessment Governing Board. *Reading framework for the National Assessment of Educational Progress: 1992–2000.*

The assessment framework specified not only the particular aspects of reading literacy to be measured, but also the percentage of the assessment questions that should be devoted to each. The target percentage distributions of reading purposes and reading stances as specified in the framework, along with the actual percentage distributions in the assessment, are presented in tables A.1 and A.2. The actual content of the assessment has varied from the targeted distribution, with Personal Response and Critical Stance Questions falling below the target proportions in the framework. The reading instrument development panel overseeing the development of the assessment recognized this variance, but felt strongly that assessment questions must be sensitive to the unique elements of the authentic reading materials being used. Thus, the distribution of question classifications will vary across reading passages and reading purposes.

The Assessment Design

Students participating in the assessment received a booklet containing a set of general background questions, reading materials and comprehension questions, reading-specific background questions, and questions about their motivation and familiarity with the assessment tasks. Reading materials that served as stimuli and their corresponding questions were assembled into sets or "blocks." Students were given two 25-minute blocks of reading passages and questions.

The grade 4 assessment consisted of eight 25-minute blocks: four blocks of literary materials and questions, and four

Table A.1

Target and actual percentage distribution of questions by reading purpose, 2000 NAEP grade 4 reading assessment

		Reading Purpose	
		Literary experience	**Gain information**
Grade 4	Target	55%	45%
	Actual	50%	50%

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Table A.2

Target and actual percentage distribution of questions by reading stance, 2000 NAEP grade 4 reading assessment

		Initial understanding/ developing an interpretation	Personal response	Critical stance
Grade 4	Target	33%	33%	33%
	Actual	59%	16%	25%

Actual percentages are based on the classifications agreed upon by NAEP's Instrument Development Panel. It is recognized that making discrete classifications for these categories is difficult and that independent efforts to classify NAEP questions have led to different results.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

blocks of informative materials and questions. Each block contained one passage corresponding to one of the reading purposes and 9 to 12 questions in both multiple-choice and constructed-response formats. In each block, one of the constructed-response questions required an extended response. As a whole, the fourth-grade assessment consisted of 35 multiple-choice questions, 38 short constructed-response questions, and 8 extended constructed-response questions.

The assessment design allowed for maximum coverage of reading abilities at grade 4, while minimizing the time burden for any one student. This was accomplished through the use of matrix sampling of items, in which representative samples of students took various portions of the entire pool of assessment questions. Individual students were required to take only a small portion of the assessment, but the aggregate results across the entire assessment allowed for broad reporting of reading abilities for the targeted population.

In addition to matrix sampling, the assessment design utilized a procedure for distributing booklets that controlled for position and context effects. Students received different blocks of passages and comprehension questions in their booklets according to a procedure called "partially balanced incomplete block (PBIB) spiraling." This procedure assigned blocks of questions so that every block appears in the first and second position within a booklet an equal number of times. Every block of questions was paired with every other block with the same reading purpose, and every block was paired with some block having the other reading purpose. The spiraling aspect of this procedure cycles the booklets for administration, so that typically only a few students in any assessment session receive the same booklet.

In addition to the student assessment booklets, three other instruments provided data relating to the assessment—a teacher questionnaire, a school questionnaire, and a Students with Disabilities/Limited English Proficiency (SD/LEP) questionnaire. The SD/LEP student questionnaire was completed by a school staff member knowledgeable about those students who were selected to participate in the assessment and who were identified as (1) having an Individualized Education Plan (IEP) or equivalent plan (for reasons other than being gifted or talented) or (2) being limited English proficient (LEP). An SD/LEP student questionnaire was completed for each identified student regardless of whether the student participated in the assessment. Each SD/LEP questionnaire took approximately three minutes to complete and asked about the student and the special programs in which he or she participated.

National Sample

The results presented in this report are based on a nationally representative probability sample of fourth-grade students. The sample was selected using a complex multistage design that involved sampling students from selected schools within selected geographic areas across the country. The sample design had the following stages:

1. selection of geographic areas (a county, group of counties, or metropolitan statistical area);
2. selection of schools (public and nonpublic) within the selected areas; and
3. selection of students within selected schools.

Each selected school that participated in the assessment and each student assessed represents a portion of the population of interest. Sampling weights are needed to make valid inferences between the student samples and the respective populations from which they were drawn. Sampling weights account for disproportionate representation due to the oversampling of students who attend schools with high concentrations of black and/or Hispanic students and students who attend nonpublic schools. Among other uses, sampling weights also account for lower sampling rates for very small schools.

A special feature of the 1998 and 2000 national assessments of reading was the collection of data from samples of students where accommodations were not permitted and samples of students where accommodations were permitted. The inclusion rules were applied, and accommodations were offered only when a student had an Individualized Education Plan (IEP) for reasons other than being gifted and talented or was identified as limited English proficient (LEP); all other students were asked to participate in the assessment.

Table A.3 shows the number of students included in the national samples for the NAEP reading assessments. For the 1998 and 2000 assessments, the table includes the number of students in the sample where accommodations were not permitted and the number of students in the sample where accommodations were permitted. The table shows that the same non-SD/LEP students were included in both samples. Only the SD/LEP students differed between the two samples.

Table A.3

Grade 4 national student sample sizes

	1992	1994	1998		2000	
	Accommodations not permitted sample	Accommodations not permitted sample	Accommodations not permitted sample	Accommodations permitted sample	Accommodations not permitted sample	Accommodations permitted sample
Non-SD/LEP students assessed	6,051	6,783	7,232		7,484	
SD/LEP students assessed without accommodations	263	599	440	413	430	476
SD/LEP students assessed who required and received accommodations	NA	NA	NA	167	NA	114
Total sample assessed	6,314	7,382	7,672	7,812	7,914	8,074

NA = Not applicable. No accommodations were permitted in this sample.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Table A.4 provides a summary of the national school and student participation rates for the reading assessment samples where accommodations were not permitted and where accommodations were permitted. Participation rates are presented for public and nonpublic schools, individually and combined. The first rate is the weighted percentage of schools participating in the assessment before substitution. This rate is based only on the number of schools that were initially selected for the assessment. The numerator of this rate is the sum of the number of students represented by each initially selected school that participated in the assessment. The denominator is the sum of the number of students represented by each of the initially selected schools that had eligible students enrolled. The initially selected schools include those that participated and those that did not.

The second school participation rate is the weighted participation rate after substitution. The numerator of this rate is the sum of the number of students represented by each of the participating schools, whether originally selected or substituted. The denominator is the same as that for the weighted participation rate for the initial sample. The denominator for this participation rate, as well as for the rate before substitution of schools, is the number of eligible students from all schools with eligible students within the nation. Because of the common denominators, the weighted participation rate after substitution is as least as great as the weighted participation rate before substitution.

Also presented in table A.4 are weighted student participation rates. The numerator of this rate is the sum across all students assessed (in either an initial session or a makeup session) of the number of students that each represents. The denominator of this rate is the sum across all eligible sampled students in participating schools of the number of students that each represents. The overall participation rates take into account the weighted percentage school participation before or after substitution and the weighted percentage student participation after makeup sessions.

Table A.4

NAEP 2000 school and student participation rates for the nation: Grade 4 public schools, nonpublic schools, and combined

	Weighted school participation			Samples where accommodations were not permitted				Samples where accommodations were permitted			
						Overall participation rate				Overall participation rate	
	Percentage before substitutes	Percentage after substitutes	Total number of schools	Weighted percentage student participation	Total number of students assessed	Before substitution	After substitution	Weighted percentage student participation	Total number of students assessed	Before substitution	After substitution
Grade 4											
Public	84	87	325	96	5,945	80	83	96	6,095	81	83
Nonpublic	86	89	108	96	1,969	83	86	97	1,979	83	86
Combined	84	87	433	96	7,914	81	84	96	8,074	81	84

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Students with Disabilities (SD) and Limited English Proficient (LEP) Students

It is NAEP's intent to assess all selected students from the target population. Therefore, every effort is made to ensure that all selected students who are capable of participating in the assessment are assessed. Some students sampled for participation in NAEP can be excluded from the sample according to carefully defined criteria. These criteria were revised in 1996 to more clearly communicate a presumption of inclusion except under special circumstances. According to these criteria, students with Individualized Education Programs (IEPs) were to be included in the NAEP assessment except in the following cases:

1. The school's IEP team determined that the student could not participate, OR,
2. The student's cognitive functioning was so severely impaired that she or he could not participate, OR,
3. The student's IEP required that the student had to be tested with an accommodation or adaptation and that the student could not demonstrate his or her knowledge without that accommodation.

All LEP students receiving academic instruction in English for three years or more were to be included in the assessment. Those LEP students receiving instruction in English for less than three years were to be included unless school staff judged them as being incapable of participating in the assessment in English.

For the 1998 and 2000 national assessments in reading, for one type of sample, the assessment was conducted using these criteria with no provisions made for accommodations. The results for this sample are presented in chapters 1, 2, and 3. For another type of sample, the assessment was conducted using these criteria, with provisions made for accommodations for identified students. These results are presented in chapter 4. The accommodations provided by NAEP were meant to match those specified in the student's IEP or those ordinarily provided in the classroom for testing situations.

Currently, NAEP is in the process of changing the way that assessments are conducted to better reflect the purpose of the 1997 Individuals with Disabilities Education Act (IDEA).[1] Permitting accommodations for identified students in the 1998 and 2000 samples laid the groundwork for future NAEP reading assessments in which the provision of accommodations will be standard program practice. Also, the NAEP 1998 and 2000 reading assessments included samples where no accommodations were permitted. These samples were comparable to samples from previous assessments so that trend results could be reported.

---

[1] Office of Special Education Programs. (1997). *Nineteenth annual report to Congress on the implementation of the individuals with disabilities education act.* Washington, DC: U.S. Department of Education.

Participation rates for students with disabilities and LEP student samples are presented in table A.5 for samples where accommodations were not permitted and in table A.6 for samples where accommodations were permitted for identified students. As can be seen from the data in table A.6, three percent of the students sampled for the 1998 and 2000 reading assessments were provided with accommodations. Ninety-seven percent of the sampled students (including SD/LEP and non-SD/LEP students), took the assessment under standard conditions.

Table A.5

Students with disabilities and limited English proficient students in the NAEP reading assessment: Grade 4 national samples where accommodations were not permitted: 1992–2000

	1992		1994		1998		2000	
	Number of students	Weighted percentage of students sampled	Number of students	Weighted percentage of students sampled	Number of students	Weighted percentage of students sampled	Number of students	Weighted percentage of students sampled
• SD and LEP students								
Identified	2,013	10	1,624	13	985	16	823	15
Excluded	1,750	6	1,025	5	545	9	393	7
Assessed	263	4	599	8	440	7	430	8
• SD students only								
Identified	1,149	7	1,039	10	490	11	524	11
Excluded	990	4	[illegible]	4	247	6	295	6
Assessed	159	3	354	6	243	5	229	5
• LEP students only								
Identified	945	3	623	4	527	6	356	5
Excluded	835	2	368	1	323	3	141	2
Assessed	110	1	255	2	204	2	215	3

SD = Students with Disabilities (the term previously used was IEP).
LEP = Limited English Proficient students.

NOTE: The combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions but counted only once in the top portion. Within each portion of the table, percentages may not sum properly due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table A.6

Students with disabilities and limited English proficient students in the NAEP reading assessment: Grade 4 national samples where accommodations were permitted for identified students: 1998 and 2000

	1998		2000	
	Number of students	Weighted percentage of students sampled	Number of students	Weighted percentage of students sampled
• SD and LEP students				
Identified	973	15	906	17
Excluded	393	6	316	6
Assessed	580	9	590	11
Assessed without accommodations	413	6	476	9
Assessed with accommodations	167	3	114	3
• SD students only				
Identified	558	11	510	12
Excluded	246	5	193	4
Assessed	312	6	317	7
Assessed without accommodations	179	4	209	5
Assessed with accommodations	133	3	108	2
• LEP students only				
Identified	446	5	446	6
Excluded	167	2	159	2
Assessed	279	3	287	4
Assessed without accommodations	238	3	273	4
Assessed with accommodations	41	1	14	▲

▲ Percentage is between 0.0 and 0.5.

SD = Students with Disabilities (the term previously used was IEP)
LEP = Limited English Proficient students

NOTE: The combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions but counted only once in the top portion. Within each portion of the table, percentages may not sum properly due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table A.7 displays the number and the percentages of SD and LEP students assessed with the variety of available accommodations. It should be noted that students assessed with accommodations typically received some combination of accommodations. For example, students assessed in small groups (as compared to standard NAEP sessions of about 30 students) usually received extended time. In one-on-one administrations, students often received assistance in recording answers and were afforded extra time. Extended time was considered the primary accommodation only when it was the sole accommodation provided. The assessment did not, however, allow some accommodations that were permitted in certain states in past assessments. Some states have allowed questions and, in some cases, reading passages to be read aloud to the students. In designing the 2000 reading assessment, reading aloud as an accommodation was viewed as changing the nature of the construct being measured and, hence, was not permitted. Because NAEP considers the domain of its reading assessment to be reading in English, no attempt was made to provide an alternate language version of the assessment, and the use of bilingual dictionaries was not permitted.

Table A.7

SD and LEP students assessed with accommodations, in the NAEP reading assessment:
Grade 4 national samples where accommodations were permitted for identified students, public and nonpublic schools combined by type of accommodation: 1998 and 2000

	1998		2000	
	Number of students	Weighted percentage of students sampled	Number of students	Weighted percentage of students sampled
SD and LEP students				
Large print	0	0	1	0.06
Extended time	63	1.07	41	0.86
Small group	90	1.94	61	1.48
One-on-one	9	0.23	9	0.27
Scribe or computer	1	0.05	1	0.03
Other	4	0.09	1	0.01
SD students only				
Large print	0	0	1	0.06
Extended time	43	0.78	41	0.86
Small group	76	1.70	55	1.36
One-on-one	9	0.23	9	0.27
Scribe or computer	1	0.05	1	0.03
Other	4	0.09	1	0.01
LEP students only				
Large print	0	0	0	0
Extended time	22	0.31	1	0.01
Small group	19	0.32	11	0.20
One-on-one	0	0	1	0.01
Scribe or computer	0	0	0	0
Other	0	0	1	0.01

SD = Students with Disabilities (the term previously used was IEP)
LEP = Limited English Proficient students

NOTE: The combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions but counted only once in each line of the top portion. Within each portion of the table, percentages may not sum properly due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Data Collection and Scoring

The 2000 reading assessment was conducted from January through March 2000, with some makeup sessions in early April. As with all NAEP assessments, data collection for the 2000 assessment was conducted by a trained field staff. For the national assessment, this was accomplished by Westat, Inc. staff.

Materials from the 2000 assessment were shipped to National Computer Systems, where trained staff evaluated the responses to the constructed-response questions using scoring rubrics or guides prepared by the Educational Testing Service. Each constructed-response question had a unique scoring rubric that defined the criteria used to evaluate students' responses. The extended constructed-response questions were evaluated with four-level rubrics, and many of the short constructed-response questions were rated according to three-level rubrics that permitted partial credit. Other short constructed-response questions were scored as either acceptable or unacceptable.

For the 2000 reading assessment, approximately 123,075 constructed responses were scored. This number includes rescoring to monitor inter-rater reliability and trend reliability. In other words, scoring reliability was calculated within year (2000) and across years (1994, 1998, and 2000). The within-year average percentage of agreement for the 2000 national grade 4 reliability sample was 88 percent. The percentages of agreement across the assessment years for the national inter-year (1994 to 2000) reliability sample was 89 percent.

Data Analysis and IRT Scaling

Subsequent to the professional scoring, all information was transcribed to the NAEP database at ETS. Each processing activity was conducted with rigorous quality control. After the assessment information had been compiled in the database, the data were weighted according to the population structure. The weighting for the national sample reflected the probability of selection for each student as a result of the sampling design, adjusted for nonresponse. Through post-stratification, the weighting assured that the representation of certain subpopulations corresponded to figures from the U.S. Census and the Current Population Survey.[2]

Analyses were then conducted to determine the percentages of students who gave various responses to each cognitive and background question. In determining these percentages for the cognitive questions, a distinction was made between missing responses at the end of a block (i.e., missing responses subsequent to the last question the student answered) and missing responses prior to the last observed response. Missing responses before the last observed response were considered intentional omissions. Missing responses at the end of the block were considered "not reached" and treated as if the questions had not been presented to the student. In calculating response percentages for each question, only students classified as having been presented the question were included in the denominator of the statistic.

---

[2] These procedures are described more fully in the section "Weighting and Variance Estimation." For additional information about the use of weighting procedures in NAEP, see Johnson, E.G. (1989, December). Considerations and techniques for the analysis of NAEP data. *Journal of Education Statistics, 14*(4), 303–334.

It is standard ETS practice to treat all nonrespondents to the last question in a block as if they had not reached the question. For multiple-choice and short constructed-response questions, this practice produces a reasonable pattern of results in that the proportion reaching the last question is not dramatically smaller than the proportion reaching the next-to-last question. However, for blocks that ended with extended constructed-response questions, the standard ETS practice would result in extremely large drops in the proportion of students attempting the final question. Therefore, for blocks ending with an extended constructed-response question, students who answered the next-to-last question but did not respond to the extended constructed-response question were classified as having intentionally omitted the last question.

Item response theory (IRT) was used to estimate average reading scale scores for the nation and for various subgroups of interest within the nation. IRT models the probability of answering a question in a certain way as a mathematical function of proficiency or skill. The main purpose of IRT analysis is to provide a common scale on which performance can be compared across groups such as those defined by characteristics, including gender and race/ethnicity.

The results for 1992, 1994, 1998, and 2000 are presented on the grade 4 NAEP reading scale. In 1992, a scale ranging from 0 to 500 was created to report performance for the literary and information reading purposes at grade 4.[3] The scales summarize student performance across all three types of questions in the assessment (multiple-choice, short constructed-response, and extended constructed-response). Results from subsequent reading assessments (1994, 1998, and 2000) are reported on these scales.

Each reading scale was initially based on the distribution of student performance across all three grades in the 1992 national assessment (grades 4, 8, and 12). In that year, the scales had an average of 250 and a standard deviation of 50. In addition, a composite scale was created as an overall measure of students' reading performance. At grade 4, this composite scale is a weighted average of the two separate scales for the two reading purposes assessed. The weight for each reading purpose is proportional to the relative importance assigned to the reading purpose by the specifications developed through the consensus planning process and given in the framework. (See target percentages in table A.1)

In producing the reading scales, three distinct IRT models were used. Multiple-choice questions were scaled using the three-parameter logistic (3PL) model; short constructed-response questions rated as acceptable or unacceptable were scaled using the two-parameter logistic (2PL) model; and short constructed-response questions rated according to a three-level rubric, as well as extended constructed-response questions rated on a four-level rubric, were scaled using a generalized partial-credit (GPC) model.[4] Developed by ETS and first used in 1992, the GPC model permits the scaling of questions scored according to multipoint rating schemes. The model takes full advantage of the information available from each of the student response categories used for these more complex constructed-response questions.

---

[3] A third purpose, reading to perform a task, was not assessed at grade 4.

[4] Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*(2), 159–176.

One natural question about the reading scales concerns the amount of information contributed by each type of question. Unfortunately, this question has no simple answer for the NAEP reading assessment, due to the complex procedures used to form the composite reading scale. The information provided by a given question is determined by the IRT model used to scale the question. It is a function of the item parameters and varies by level of reading proficiency.[5] Thus, the answer to the query "How much information do the different types of questions provide?" will differ for each level of reading performance. When considering the composite reading scale, the answer is even more complicated. The reading data are scaled separately by the purposes of reading (reading for literary experience, reading to gain information, and reading to perform a task).[6] The composite scale is a weighted combination of these subscales. IRT information functions are only strictly comparable when they are derived from the same calibration. Because the composite scale is based on two separate calibrations, there is no direct way to compare the information provided by the questions on the composite scale.

Because of the PBIB-spiraling design used by NAEP, students do not receive enough questions about a specific topic to provide reliable information about individual performance. Traditional test scores for individual students, even those based on IRT, would lead to misleading estimates of population characteristics, such as subgroup means and percentages of students at or above a certain scale-score level. Consequently, NAEP constructs sets of plausible values designed to represent the distribution of performance in the population. A plausible value for an individual is not a scale score for that individual, but may be regarded as a representative value from the distribution of potential scale scores for all students in the population with similar characteristics and identical patterns of item response. Statistics describing performance on the NAEP reading scale are based on the plausible values. Under the assumptions of the scaling models, these population estimates will be consistent, in the sense that the estimates approach the model-based population values as the sample size increases, which would not be the case for population estimates obtained by aggregating optimal estimates of individual performance.[7]

---

[5] Donoghue, J.R. (1994). An empirical examination of the IRT information of polytomously scored reading items under the generalized partial credit model. *Journal of Educational Measurement, 31*(4), 295–311.

[6] Only two purposes, literary and information, were used in the 2000 grade 4 reading assessment. These are described in the introduction of this report.

[7] For theoretical and empirical justification of the procedures employed, see Mislevy, R.J. (1988). Randomization-based inferences about latent variables from complex samples. *Psychometrika, 56*(2), 177–196.

For computational details, see the forthcoming NAEP 2000 technical report. National Assessment of Educational Progress. (2000). *NAEP 2000 technical report.* [forthcoming] Princeton, NJ: Educational Testing Service.

Item Mapping Procedures

To map items to particular points on the reading proficiency scale, a response probability convention had to be adopted that would divide those who had a higher probability of success from those who had a lower probability. Establishing a response probability convention has an impact on the mapping of the test items onto the reading scale. A lower boundary convention maps the reading items at lower points along the scale, and a higher boundary convention maps the same items at higher points on the scale. The underlying distribution of reading skills in the population does not change, but the choice of a response probability convention does have an impact on the proportion of the student population that is reported as "able to do" the items on the reading scales.

There is no obvious choice of a point along the probability scale that is clearly superior to any other point. If the convention were set with a boundary at 50 percent, those above the boundary would be more likely to get an item right than get it wrong, while those below the boundary would be more likely to get the item wrong than right. Although this convention has some intuitive appeal, it was rejected on the grounds that having a 50/50 chance of getting the item right shows an insufficient degree of mastery. If the convention were set with a boundary at 80 percent, students above the criterion would have a high probability of success with an item. However, many students below this criterion show some level of reading ability that would be ignored by such a stringent criterion. In particular, those in the range between 50 and 80 percent correct would be more likely to get the item right than wrong, yet would not be in the group described as "able to do" the item.

In a compromise between the 50 percent and the 80 percent conventions, NAEP has adopted two related response probability conventions: 74 percent for multiple-choice questions (to correct for the possibility of answering correctly by guessing) and 65 percent for constructed-response questions (where guessing is not a factor). These probability conventions were established, in part, based on an intuitive judgment that they would provide the best picture of students' reading skills.

Some additional support for the dual conventions adopted by NAEP was provided by Huynh.[8] He examined the IRT information provided by items, according to the IRT model used in scaling NAEP questions. ("Information" is used here in a technical sense. See the forthcoming *NAEP 2000 Technical Report* for details.) Following Bock,[9] Huynh decomposed the item information into that provided by a correct response [P(q) I(q)] and that provided by an incorrect response [(1- P(q)) I(q)]. Huynh showed that the item information provided by a correct response to a constructed-response item is maximized at the point along the reading scale at which the probability of a correct response is two thirds (for multiple-choice items, the information provided by a correct response is maximized at the point at which the probability of getting the item correct is .74). It should be noted, however, that maximizing the item information I(q), rather than the information provided by a correct response [P(q) I(q)], would imply an item mapping criterion closer to 50 percent.

---

[8] Huynh, H. (1994, October). *Some technical aspects of standard setting.* Paper presented at the Joint Conference on Standard Setting for Large-Scale Assessment, Washington, DC.

[9] Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more latent categories. *Psychometrika*, 37, 29–51.

The results in this report are presented in terms of the composite reading scale. However, the reading assessment was scaled separately for the two purposes for reading at grade 4. The composite scale is a weighted combination of the two subscales for purposes for reading. To obtain item map information presented in this report, a procedure by Donoghue was used.[39] This method models the relationship between the item response function for the subscale and the subscale structure to derive the relationship between the item score and the composite scale (i.e., an item response function for the composite scale). This item response function is then used to derive the probability used in the mapping.

Weighting and Variance Estimation

A complex sample design was used to select the students who were assessed. The properties of a sample selected through a complex design could be very different from those of a simple random sample, in which every student in the target population has an equal chance of selection and in which the observations from different sampled students can be considered to be statistically independent of one another. Therefore, the properties of the sample for the complex data collection design were taken into account during the analysis of the assessment data.

One way that the properties of the sample design were addressed was by using sampling weights to account for the fact that the probabilities of selection were not identical for all students. All population and subpopulation characteristics based on the assessment data used sampling weights in their estimation. These weights included adjustments for school and student nonresponse.

Not only must appropriate estimates of population characteristics be derived, but appropriate measures of the degree of uncertainty must be obtained for those statistics. Two components of uncertainty are accounted for in the variability of statistics based on student ability: (1) the uncertainty due to sampling only a relatively small number of students, and (2) the uncertainty due to sampling only a relatively small number of cognitive questions. The first component accounts for the variability associated with the estimated percentages of students who had certain background characteristics or who answered a certain cognitive question correctly.

Because NAEP uses complex sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate. NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any student information that can be observed without error. However, because each student typically responds to only a few questions within any purpose for reading, the scale score for any single student would be imprecise. In this case, plausible values methodology can be used to describe the performance of groups and subgroups of students, but the underlying imprecision involved in this step adds another component of variability to statistics based on NAEP scale scores.[11] (Appendix C provides the standard errors for the results presented in this report.)

---

[39] Donoghue, J. R. (1997, March). *Item mapping to a weighted composite scale.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Typically, when the standard error is based on a small number of students or when the group of students is enrolled in a small number of schools, the amount of uncertainty associated with the standard errors may be quite large. Throughout this report, estimates of standard errors subject to a large degree of uncertainty are followed by the "!" symbol. In such cases, the standard errors—and any confidence intervals or significance tests involving these standard errors—should be interpreted cautiously. Additional details concerning procedures for identifying such standard errors are discussed in the forthcoming *NAEP 2000 Technical Report*.

The reader is reminded that, as with findings from all surveys, NAEP results are subject to other kinds of error, including the effects of imperfect adjustment for student and school nonresponse and unknowable effects associated with the particular instrumentation and data collection methods. Nonsampling errors can be attributed to a number of sources—inability to obtain complete information about all selected schools in the sample (some students or schools refused to participate, or students participated but answered only certain questions); ambiguous definitions; differences in interpreting questions; inability or unwillingness to give correct information; mistakes in recording, coding, or scoring data; and other errors in collecting, processing, sampling, and estimating missing data. The extent of nonsampling error is difficult to estimate; and, because of their nature, the impact of such errors cannot be reflected in the data-based estimates of uncertainty provided in NAEP reports.

Drawing Inferences from the Results

Because the percentages of students in these subpopulations and their average scale scores are based on samples rather than on the entire population of fourth-graders in the nation, the numbers reported are estimates. As such, they are subject to a measure of uncertainty, reflected in the standard error of the estimate. When the percentages or average scale scores of certain groups are compared, the standard error should be taken into account, and observed similarities or differences should not be relied on solely. Therefore, the comparisons discussed in this report are based on statistical tests that consider the standard errors of those statistics and the magnitude of the difference among the averages or percentages.

Using confidence intervals based on the standard errors provides a way to take into account the uncertainty associated with sample estimates, and to make inferences about the population averages and percentages in a manner that reflects that uncertainty. An estimated sample average scale score plus or minus two standard errors approximates a 95 percent confidence interval for the corresponding population quantity. This statement means that one can conclude with approximately a 95 percent level of confidence that the average performance of the entire population of interest (e.g., all fourth-grade students in public and nonpublic schools) is within plus or minus two standard errors of the sample average.

---

[11] For further details, see Johnson, E.G. & Rust, K.F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics, 17*(2), 175–190.

As an example, suppose that the average reading scale score of the students in a particular group was 256 with a standard error of 1.2. A 95 percent confidence interval for the population quantity would be as follows:

$$\text{Average} \pm 2 \text{ standard errors}$$

$$256 \pm 2 \times 1.2$$

$$256 \pm 2.4$$

$$(253.6, 258.4)$$

Thus, one can conclude with a 95 percent level of confidence that the average scale score for the entire population of students in that group is between 253.6 and 258.4.

Similar confidence intervals can be constructed for percentages, if the percentages are not extremely large or extremely small. Extreme percentages should be interpreted with caution. Adding or subtracting the standard errors associated with extreme percentages could cause the confidence interval to exceed 100 percent or go below 0 percent, resulting in numbers that are not meaningful. (The forthcoming *NAEP 2000 Technical Report* will contain a more complete discussion of extreme percentages.)

Analyzing Group Differences in Averages and Percentages

The statistical tests determine whether the evidence, based on the data from the groups in the sample, is strong enough to conclude that the averages or percentages are actually different for those groups in the population. If the evidence is strong (i.e., the difference is statistically significant), the report describes the group averages or percentages as being different (e.g., one group performed higher than or lower than another group), regardless of whether the sample averages or percentages appear to be approximately the same. Occasionally, if an apparent difference is quite large but not statistically significant, this report will point out that fact.

The reader is cautioned to rely on the results of the statistical tests rather than on the apparent magnitude of the difference between sample averages or percentages when determining whether the sample differences are likely to represent actual differences among the groups in the population.

To determine whether a real difference exists between the average scale scores (or percentages of a certain attribute) for two groups in the population, one needs to obtain an estimate of the degree of uncertainty associated with the difference between the averages (or percentages) of these groups for the sample. This estimate of the degree of uncertainty, called the standard error of the difference between the groups, is obtained by taking the square of each group's standard error, summing the squared standard errors, and taking the square root of that sum.

Standard Error of the Difference =

$$SE_{A-B} = \sqrt{(SE_A^2 + SE_B^2)}$$

Similar to how the standard error for an individual group average or percentage is used, the standard error of the difference can be used to help determine whether differences among groups in the population are real. The difference between the averages or percentages of the two groups plus or minus two standard errors of the difference represents an approximate 95 percent confidence interval. If the resulting interval includes zero, there is insufficient evidence to claim a real difference between the groups in the population. If the interval does not contain zero, the difference between the groups is statistically significant (different) at the 0.05 level.

As an example of comparing groups, consider the problem of determining whether the average reading scale score of group A is higher than that of group B. Suppose that the sample estimates of the average scale scores and standard errors were as follows:

Group	Average Scale Score	Standard Error
A	218	0.9
B	216	1.1

The difference between the estimates of the average scale scores of groups A and B is two points (218 − 216). The standard error of this difference is

$$\sqrt{(0.9^2 + 1.1^2)} = 1.4$$

Thus, an approximate 95 percent confidence interval for this difference is plus or minus two standard errors of the difference

$$2 \pm 2 \times 1.4$$

$$2 \pm 2.8$$

$$(-0.8, 4.8)$$

The value zero is within the confidence interval; therefore, there is insufficient evidence to claim that group A outperformed group B.

In some cases, the differences between groups were not discussed in this report. This happened for one of two reasons: (a) if the comparison involved an extreme percentage (as defined above); or (b) if the standard error for either group was subject to a large degree of uncertainty (i.e., the coefficient of variation is greater than 20 percent, denoted by "!" in the tables).[12] In either case, the results of any statistical test involving that group need to be interpreted with caution; and so, the results of such tests are not discussed in this report.

---

[12] As was discussed in the section "Weighting and Variance Estimation," estimates of standard errors subject to a large degree of uncertainty are designated by the symbol "!". In such cases, the standard error—and any confidence intervals or significance tests among these standard errors—should be interpreted with caution.

Conducting Multiple Tests

The procedures in the previous section and the certainty ascribed to intervals (e.g., a 95 percent confidence interval) are based on statistical theory that assumes that only one confidence interval or test of statistical significance is being performed. However, in chapters 2 and 3 of this report, many different groups are being compared (i.e., multiple sets of confidence intervals are being analyzed). In sets of confidence intervals, statistical theory indicates that the certainty associated with the entire set of intervals is less than that attributable to each individual comparison from the set. To hold the significance level for the set of comparisons at a particular level (e.g., 0.05), adjustments (called "multiple comparison procedures"[13]) must be made to the methods described in the previous section. One such procedure, the False Discovery Rate (FDR) procedure[14] was used to control the certainty level.

Unlike the other multiple comparison procedures (e.g., the Bonferroni procedure) that control the familywise error rate (i.e., the probability of making even one false rejection in the set of comparisons), the FDR procedure controls the expected proportion of falsely rejected hypotheses. Furthermore, familywise procedures are considered conservative for large families of comparisons.[15] Therefore, the FDR procedure is more suitable for multiple comparisons in NAEP than other procedures. A detailed description of the FDR procedure appears in the forthcoming *NAEP 2000 Technical Report.*

To illustrate how the FDR procedure is used, consider the comparisons of 2000 and 1994 grade 4 average reading scale scores for all five of the race/ethnicity groups presented in table A.8. Note that the

Table A.8

FDR comparisons of average reading scale scores for race/ethnicity groups: 1994 and 2000

	1994		2000					
	Average scale score	Standard error	Average scale score	Standard error	Difference in averages	Standard error of difference	Test statistic	Percent confidence
White	224	1.3	226	1.0	2.08	1.62	1.29	20
Black	187	1.7	193	1.7	6.31	2.36	2.68	1
Hispanic	191	2.6	197	1.7	6.63	3.08	2.15	4
Asian/Pacific Islander	229	4.4	232	4.6	3.24	6.35	.51	62
American Indian	201	3.4	196	4.7	-5.51	5.81	-.95	35

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

---

[13] Miller, R.G. (1966). *Simultaneous statistical inference.* New York: Wiley.

[14] Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, No. 1.,* pp 298–300.

[15] Williams, V.S.L., Jones, L.V., & Tukey, J.W. (1994, December). *Controlling error in multiple comparisons with special attention to the National Assessment of Educational Progress.* Research Triangle Park, NC: National Institute of Statistical Sciences.

difference in average scale scores and the standard error of the difference are calculated in a way comparable with that of the example in the previous section. The test statistic shown is the difference in average scale scores divided by the standard error of the difference.

The difference in average scale scores and its standard error can be used to find an approximate 95 percent confidence interval as in the example in the previous section or they can be used to identify a confidence percentage. In the example in the previous section, because an approximate 95 percent confidence interval was desired, the number 2 was used to multiply the standard error of the difference to create the approximate confidence interval. In the current example, the test statistic is treated like the number 2 and the matching percent confidence for the related confidence interval is identified from statistical tables. Instead of checking to see if zero is within the 95 percent confidence interval, the percent confidence from the statistical tables can be directly compared to 100-95 = 5 percent.

If the comparison of average scale scores across two years were made for only one of the race/ethnicity groups, there would be a significant difference between the average scale scores for the two years if the percent confidence were less than 5 percent. However, because we are interested in the difference in average scale scores across the two years for all five of the race/ethnicity groups, comparing each of the percents of confidence to 5 percent is not adequate. In this report, comparisons of average scale score differences across years for all race/ethnicity groups are discussed together. Groups of students defined by shared characteristics, such as race/ethnicity groups, are treated as sets or families when making comparisons. However, comparisons of average scale scores for each pair of years were treated separately. So the steps described in this example were replicated for the comparison of 2000 and 1998 average scale scores and the comparison of 2000 and 1992 average scale scores.

To use the FDR procedure to take into account that all five race/ethnicity comparisons are of interest to us, the percents of confidence in the example are ordered from largest to smallest: 62, 35, 20, 4, and 1. In the FDR procedure, 62 percent confidence for the Asian/Pacific Islander comparison would be compared to 5 percent, 35 percent for the American Indian comparison would be compared to 5*(5-1)/5 = 4 percent, 20 percent for the White comparison would be compared to 5*(5-2)/5 = 3 percent, 4 percent for the Hispanic comparison would be compared to 5*(5-3)/5 = 2 percent, and 1 (actually slightly smaller than 1 prior to rounding) percent for the black comparison would be compared to 5*(5-4)/5 = 1 percent. The last of these comparisons is the only one for which the percent confidence is smaller than the FDR procedure value. The difference in 2000 and 1994 average scale scores for the black students is significant; for all of the other race/ethnicity groups, average scale scores for 2000 and 1994 are not significantly different from one another.

NAEP Reporting Groups

In this report, results are provided for groups of students defined by shared characteristics—region of the country, gender, race or ethnicity, school's type of location, eligibility for the Free/Reduced-Price School Lunch program, and type of school. Based on participation rate criteria, results are reported for subpopulations only when sufficient numbers of students and adequate school representation are present. The minimum requirement is at least 62 students in a particular subgroup from at least five primary sampling units (PSUs).[16] However, the data for all students, regardless of whether their subgroup was reported separately, were included in computing overall results. Definitions of the subpopulations referred to in this report are presented below:

Figure A.2

States included in the four NAEP regions

Northeast	Southeast	Central	West
Connecticut	Alabama	Illinois	Alaska
Delaware	Arkansas	Indiana	Arizona
District of Columbia	Florida	Iowa	California
Maine	Georgia	Kansas	Colorado
Maryland	Kentucky	Michigan	Hawaii
Massachusetts	Louisiana	Minnesota	Idaho
New Hampshire	Mississippi	Missouri	Montana
New Jersey	North Carolina	Nebraska	Nevada
New York	South Carolina	North Dakota	New Mexico
Pennsylvania	Tennessee	Ohio	Oklahoma
Rhode Island	*Virginia	South Dakota	Oregon
Vermont	West Virginia	Wisconsin	Texas
*Virginia			Utah
			Washington
			Wyoming

* NOTE: The part of Virginia that is included in the Washington, DC, metropolitan area is included in the Northeast region; the remainder of the state is included in the Southeast region.

---

[16] For the national assessment, a PSU is a selected geographic region (a county, group of counties, or metropolitan statistical area). For the state assessment program, a PSU is most often a single school. Further details about the procedure for determining minimum sample size appear in the *1998 NAEP Technical Report*.

Region

Results in NAEP are reported for four regions of the nation: Northeast, Southeast, Central, and West. Figure A.2 shows how states are subdivided into these NAEP regions. All 50 states and the District of Columbia are listed. Territories and the two Department of Defense Educational Activities jurisdictions are not assigned to any region.

Gender

Results are reported separately for males and females.

Race/Ethnicity

The race/ethnicity variable is derived from two questions asked of students and from school records, and it is used for race/ethnicity subgroup comparisons. Two questions from the set of general student background questions were used to determine race/ethnicity:

If you are Hispanic, what is your Hispanic background?

☐ I am not Hispanic
☐ Mexican, Mexican American, or Chicano
☐ Puerto Rican
☐ Cuban
☐ Other Spanish or Hispanic background

Students who responded to this question by filling in the second, third, fourth, or fifth oval were considered Hispanic. For students who filled in the first oval, did not respond to the question, or provided information that was illegible or could not be classified, responses to the following question were examined to determine their race/ethnicity.

Which best describes you?

☐ White (not Hispanic)
☐ Black (not Hispanic)
☐ Hispanic ("Hispanic" means someone who is Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or other Spanish or Hispanic background)
☐ Asian or Pacific Islander ("Asian or Pacific Islander" means someone who is from a Chinese, Japanese, Korean, Filipino, Vietnamese, Asian American or from some other Asian or Pacific Islander background.)
☐ American Indian or Alaskan Native ("American Indian or Alaskan Native" means someone who is from one of the American Indian tribes or one of the original people of Alaska.)
☐ Other (specify) ____________________

Students' race/ethnicity was then assigned on the basis of their responses. For students who filled in the sixth oval ("Other"), provided illegible information or information that could not be classified, or did not respond at all, race/ethnicity was assigned as determined by school records.

Race/ethnicity could not be determined for students who did not respond to either of the demographic questions and whose schools did not provide information about race/ethnicity.

Details of how race/ethnicity classifications were derived are presented so that readers can determine how useful the results are for their particular purposes. Also, some students indicated that they were from a Hispanic background (e.g., Puerto Rican or Cuban) and that a racial/ethnic category other than Hispanic best described them. These students were classified as Hispanic based on the rules described above. Furthermore, information from the schools did not always correspond to how students described themselves.

Therefore, the racial/ethnic results presented in this report attempt to provide a clear picture based on several sources of information.

In the 1992, 1998 and 2000 NAEP reading assessments the mutually exclusive racial/ethnic categories were: white, black, Hispanic, Asian/Pacific Islander, and American Indian (including Alaskan native). In the 1994 NAEP reading assessment, the Asian Pacific/Islander category was divided into separate Asian and Pacific Islander categories. To make comparisons of performance across all three assessments, the separate Asian and Pacific Islander categories used in 1994 have been collapsed into a single category to report results.

Type of Location

Results from the 2000 assessment are reported for students attending schools in three mutually exclusive location types: central city, urban fringe/large town, and rural/small town:

Central City: This category includes central cities of all Standard Metropolitan Statistical Areas (SMSA) as defined by the Office of Management and Budget. Central City is a geographical term and is not synonymous with "inner city."

Urban Fringe/Large Town: The urban fringe category includes all densely settled places and areas within SMSA's that are classified as urban by the Bureau of the Census, but which do not qualify as Central City. A Large Town is defined as a place outside a SMSA with a population greater than or equal to 25,000.

Rural/Small Town: Rural includes all places and areas with populations of less than 2,500 that are classified as rural by the Bureau of the Census. A Small Town is defined as a place outside a SMSA with a population of less than 25,000, but greater than or equal to 2,500.

In this report, results for each type of location are not compared across years. This was due to new methods used by NCES to identify the type of location assigned to each school in the Common Core of Data (CCD). The new methods were put into place by NCES in order to improve the quality of the assignments and they take into account more information about the exact physical location of the school.

Eligibility for the Free/Reduced-Price School Lunch Program

Based on available school records, students were classified as either currently eligible for the free/reduced-price lunch component of the Department of Agriculture's National School Lunch Program or not eligible. The classification applies only to the school year when the assessment was administered (i.e., the 1999–2000 school year) and is not based on eligibility in previous years. If school records were not available, the student was classified as "Information not available." If the school did not participate in the program, all students in that school were classified as "Information not available."

Type of School

Results are reported by the type of school that the student attends—public or nonpublic. Nonpublic schools include Catholic and other private schools. Although Bureau of Indian Affairs (BIA) schools and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) are not included in either the public or nonpublic categories, they are included in the overall national results.

Cautions in Interpretations

As described earlier, the NAEP reading scale makes it possible to examine relationships between students' performance and various background factors measured by NAEP. However, a relationship that exists between achievement and another variable does not reveal its underlying cause, which may be influenced by a number of other variables. Similarly, the assessments do not capture the influence of unmeasured variables. The results are most useful when they are considered in combination with other knowledge about the student population and the educational system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.

B

Appendix B

Sample Text and Questions from the NAEP 2000 Reading Assessment

This appendix contains the reading passage released from the NAEP 2000 reading assessment. In addition, sample questions are presented to supplement those in chapter 1. The tables accompanying each question present two types of percentages: the overall percentage of students who answered successfully, and the percentage of students who answered successfully within the achievement level score ranges. For each question, the reading purpose and reading stance is indicated. For multiple-choice questions, the correct answer is marked. For constructed-response questions, the rating assigned to a response and a summary of the scoring criteria used to rate responses is provided. Sample student responses have been reproduced from actual test booklets to illustrate answers that demonstrated at least adequate comprehension. To review additional passages and questions from previous NAEP assessments, please visit NAEP on the internet at http://nces.ed.gov/nationsreportcard.

Appendix Focus

Released materials from the 2000 reading assessment

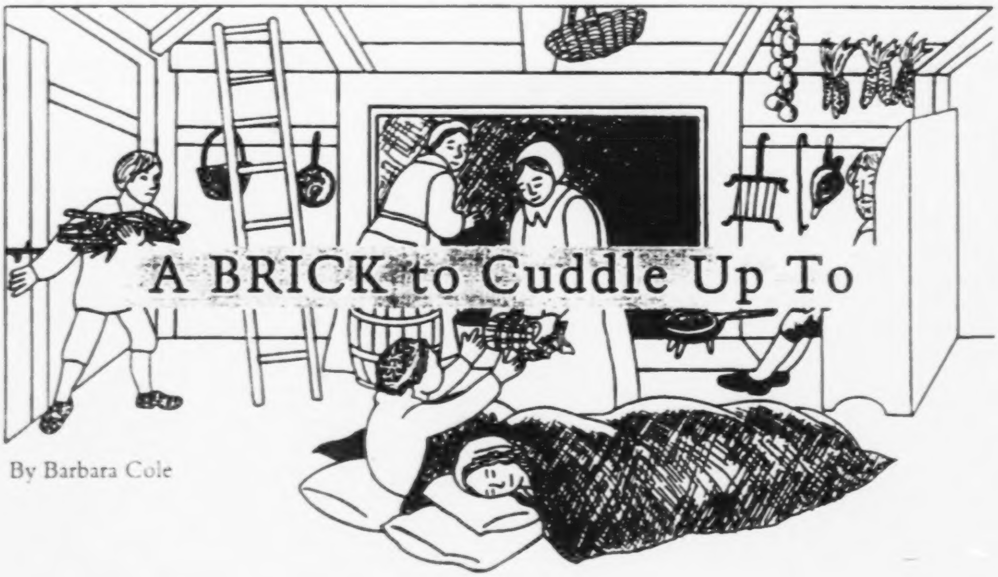
Appendix Contents

Sample Text

Sample Questions

Student Responses

A BRICK to Cuddle Up To



By Barbara Cole

Imagine shivering on a cold winter's night. The tip of your nose tingles in the frost air. Finally, you climb into bed and find the toasty treat you have been waiting for — your very own hot brick.

If you had lived in colonial days, that would not sound as strange as it does today. Winters were hard in this New World, and the colonists had to think of clever ways to fight the cold. At bedtime, they heated soapstones, or bricks, in the fireplaces. They wrapped the bricks in cloths and tucked them into their beds. The brick kept them warm at night, at least for as long as its heat lasted.

Before the colonists slipped into bed, they rubbed their icy sheets with a bed warmer. This was a metal pan with a long wooden handle. The pan held hot embers from the fireplace. It warmed the bedding so well that sleepy bodies had to wait until the sheets cooled before climbing in.

Staying warm wasn't just a bedtime problem. On winter rides, colonial travelers covered themselves with animal skins and warm blankets. Tucked under the blankets, near their feet, were small tin boxes called foot stoves. A foot stove held burning coals. Hot smoke puffed from small holes in the stove's lid, soothing freezing feet and legs. When the colonists went to Sunday services, their foot stoves, furs, and blankets went with them. The meeting houses had no heat of their own until the 1800s.

At home, colonist families huddled close to the fireplace, or hearth. The fireplace was wide and high enough to hold a large fire, but its chimney was large, too. That caused a problem: Gusts of cold air blew into the house. The area near the fire was warm, but in the rest of the room it might still be cold enough to see your breath.

Reading or needlework was done by candlelight, or by the light of the fire. During the winter, animal skins sealed the drafty windows of some cabins and blocked out the daylight. The living area inside was gloomy, except in the circle of light at the hearth.

Early Americans did not bathe as often as we do. When they did, their "bathroom" was the kitchen, in that toasty space by the hearth. They partially filled a tub of cold water, then warmed it up with water heated in the fireplace. A blanket draped from chairs for privacy also let the fire's warmth surrounded the bather.

The household cooks spent hours at the hearth. They stirred the kettle of corn pudding or checked the baking bread while the rest of the family carried on their own fireside activities. So you can see why the fireplace was the center of a colonial home.

The only time the fire was allowed to die down was at bedtime. Ashes would be piled over the fire, reducing it to embers that might glow until morning.

By sunrise, the hot brick had become a cold stone once more. An early riser might get dressed under the covers, then hurry to the hearth to warm up.

Maybe you'd enjoy hearing someone who kept warm in these ways tell you what it was like. You wouldn't need to look for someone who has been living for two hundred years. In many parts of the country the modern ways didn't take over from the old ones until recently. Your own grandparents or other older people might remember the warmth of a hearthside and the joy of having a brick to cuddle up to.

Used by permission of Highlights for Children, Inc., Columbus, OH. Copyright © 1991. Illustrations by Katherine Dodge.

Sample Question 4:

You would probably read this article if you wanted to know how the colonists

Ⓐ cooked their food

Ⓑ traveled in the winter

Ⓒ washed their clothes

● kept warm in cold weather

Reading Purpose:
To Gain Information

Reading Stance:
Initial Understanding

Table B1: Sample Question 4 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2000

	Percentage correct within achievement level intervals		
Overall percentage correct	***Basic*** 208–237*	***Proficient*** 238–267*	***Advanced*** 268 and above*
85 (0.8)	91 (2.0)	97 (1.1)	100 (****)

Includes fourth-grade students who were below the *Basic* level.
*NAEP Reading composite scale range.
Standard errors of the estimated percentages appear in parentheses.
(****) Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Sample Question 5:

Give two reasons stated in the article why the hearth was the center of the home in colonial times.

Reading Purpose:
To Gain Information

Reading Stance:
Developing an Interpretation

Responses to this question were scored according to a three-level rubric as Unsatisfactory, Partial, or Complete

Table B2: Sample Question 5 Results (Short Constructed-Response)

Overall percentage "Complete" and percentages "Complete" within each achievement-level range: 2000

	Percentage "Complete" within achievement level intervals		
Overall percentage "Complete" ¹	***Basic* 208–237***	***Proficient* 238–267***	***Advanced* 268 and above***
20 (1.2)	16 (1.8)	37 (3.2)	58 (4.9)

¹ Includes fourth-grade students who were below the *Basic* level.
*NAEP Reading composite scale range.
Standard errors of the estimated percentages appear in parentheses.
(****) Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Responses scored "Complete" demonstrated understanding of a major aspect of the article by providing two reasons for the hearth being the center of the colonial home.

Sample "Complete" Response:

Give two reasons stated in the article why the hearth was the center of the home in colonial times.

The hearth was the center of the home becaus it kept you warm and you could see by it's light.

BEST COPY AVAILABLE

Sample Question 6:

Does the author help you understand what colonial life was like? Use examples from the article to explain why or why not.

Reading Purpose:
To Gain Information

Reading Stance:
Critical Stance

Responses to this question were scored according to a three-level rubric as Unsatisfactory, Partial, or Complete

Table B3: Sample Question 6 Results (Short Constructed-Response)

Overall percentage "Complete" and percentages "Complete" within each achievement level range: 2000

	Percentage "Complete" within achievement level intervals		
Overall percentage "Complete"[1]	***Basic*** **208–237***	***Proficient*** **238–267***	***Advanced*** **268 and above***
20 (1.4)	19 (3.1)	29 (3.6)	35 (6.2)

[1] Includes fourth-grade students who were below the *Basic* level.
*NAEP Reading composite scale range.
Standard errors of the estimated percentages appear in parentheses.
(****) Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Responses scored "Complete" provided an opinion of the author's presentation of information and an example from the text to illustrate that opinion.

Sample "Complete" Response:

Does the author help you understand what colonial life was like? Use examples from the article to explain why or why not.

Yes, She does. She gives specific details a how they did the thing: they did to keep worm. She gave the specific details of how a food stove was used. And and when it was used.

Sample Question 7:

Some of the ways that colonists kept warm during the winter were different from the ways that people keep warm today. Tell about two of these differences.

Reading Purpose:
To Gain Information

Reading Stance:
Personal Reflection and Response

Responses to this question were scored according to a three-level rubric as Unsatisfactory, Partial, or Complete

Table B4: Sample Question 7 Results (Short Constructed-Response)

Overall percentage "Complete" and percentages "Complete" within each achievement level range: 2000

	Percentage "Complete" within achievement level intervals		
Overall percentage "Complete" [1]	***Basic*** **208–237***	***Proficient*** **238–267***	***Advanced*** **268 and above***
17 (1.1)	15 (2.4)	27 (3.7)	46 (6.5)

[1] Includes fourth-grade students who were below the *Basic* level.
*NAEP Reading composite scale range.
Standard errors of the estimated percentages appear in parentheses.
(****) Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Responses scored "Complete" connected text descriptions to prior knowledge by comparing two ways colonists kept warm during winter to the ways people keep warm today.

Sample "Complete" Response:

Some of the ways that colonists kept warm during the winter were different from the ways that people keep warm today. Tell about two of these differences.

Back in the "old days" they had fireplaces and bed warmers now we have heaters and electric blankets.

C

Appendix C
Data Appendix

This appendix contains complete data for all the tables and figures presented in this report, including average scores, achievement level results, and percentages of students. In addition, standard errors appear in parentheses next to each scale score and percentage. The comparisons presented in this report are based on statistical tests that consider the magnitude of the difference between group averages or percentages and the standard errors of those statistics. Because NAEP scores and percentages are based on samples rather than the entire population(s), the results are subject to a measure of uncertainty reflected in the standard errors of the estimates. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. As with the figures and tables in the chapters, significant differences between results of previous assessments and the 2000 assessment are highlighted.

Appendix Focus

Complete data for all tables and figures.

Appendix Contents

Average Scores

Achievement Level Results

Percentages of Students

Standard Errors

Table C.1: Data for Figure 1.1 Scale Score Results

Average fourth-grade reading scale scores: 1992–2000

1992	217 (0.9)
1994	214 (1.0)
1998	217 (0.8)
2000	217 (0.8)

Standard errors of the estimated scale scores appear in parentheses.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.2: Data for Figure 1.2 Performance Distribution

Fourth-grade reading scale score percentiles: 1992–2000

	Mean	10th	25th	50th	75th	90th
1992	217 (0.9)	170 (1.9) *	194 (1.1)	219 (1.3)	242 (1.1) *	261 (1.4) *
1994	214 (1.0)	159 (1.5) *	189 (1.0) *	219 (1.2)	243 (1.3)	263 (1.6)
1998	217 (0.8)	167 (1.5)	193 (0.9)	220 (1.3)	244 (0.9)	263 (0.9)
2000	217 (0.8)	163 (1.9)	193 (0.9)	221 (1.1)	245 (0.8)	264 (0.9)

Standard errors of the estimated scale scores appear in parentheses.
* Significantly different from 2000.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.3: Data for Figure 1.4 Achievement Level Results

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range: 1992–2000

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
1992	38 (1.1)	34 (0.9) *	22 (0.9)	6 (0.6) *	62 (1.1)	29 (1.2) *
1994	40 (1.0)	31 (0.7)	22 (0.8)	7 (0.7)	60 (1.0)	30 (1.1)
1998	38 (0.9)	32 (0.7)	24 (0.7)	7 (0.5)	62 (0.9)	31 (0.9)
2000	37 (0.8)	31 (0.9)	24 (0.8)	8 (0.5)	63 (0.8)	32 (0.9)

Standard errors of the estimated percentages appear in parentheses.

* Significantly different from 2000.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.4: Data for Table 1.1 Sample Question 1 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2000

	Overall	*Basic*	*Proficient*	*Advanced*
2000	66 (1.6)	72 (3.5)	79 (3.3)	84 (5.1)

Standard errors of the estimated percentages appear in parentheses.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Table C.5: Data for Table 1.2 Sample Question 2 Results (Short Constructed-Response)

Overall percentage "Complete" and percentages "Complete" within each achievement level range: 2000

	Overall	*Basic*	*Proficient*	*Advanced*
2000	37 (1.6)	38 (3.6)	57 (3.9)	76 (4.1)

Standard errors of the estimated percentages appear in parentheses.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Table C.6: Data for Table 1.3 Sample Question 3 Results (Extended Constructed-Response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement level range: 2000

	Overall	*Basic*	*Proficient*	*Advanced*
2000	18 (1.0)	15 (1.9)	29 (3.2)	40 (4.9)

Standard errors of the estimated percentages appear in parentheses.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Table C.7: Data for Figure 2.1 Scale Score Results by Gender

Percentage of students and average fourth-grade reading scale scores for male and female students: 1992–2000

	Male	Female
1992	51 (0.6) 213 (1.2)	49 (0.6) 221 (1.0)
1994	51 (0.7) 209 (1.3)	49 (0.7) 220 (1.1)
1998	50 (0.6) 214 (1.1)	50 (0.6) 220 (0.7)
2000	50 (0.7) 212 (1.1)	50 (0.7) 222 (0.9)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.8: Data for Figure 2.2 Achievement Level Results by Gender

Percentages of fourth-grade male and female students at or above reading achievement levels and within each achievement level range: 1992–2000

		Below Basic	*At Basic*	*At Proficient*	*Advanced*	*At or above Basic*	*At or above Proficient*
Male	1992	42 (1.6)	32 (1.3)	20 (1.1)	5 (0.7)	58 (1.6)	25 (1.4)
	1994	45 (1.4)	30 (1.1)	20 (0.9)	6 (0.8)	55 (1.4)	26 (1.3)
	1998	41 (1.4)	31 (1.0)	22 (0.9)	6 (0.6)	59 (1.4)	28 (1.2)
	2000	42 (1.2)	31 (1.1)	21 (1.2)	6 (0.5)	58 (1.2)	27 (1.1)
Female	1992	33 (1.3)	35 (1.4)	24 (1.2)	8 (0.8)	67 (1.3)	32 (1.4) *
	1994	34 (1.2)	32 (1.1)	25 (1.5)	9 (0.9)	66 (1.2)	34 (1.5)
	1998	35 (1.0)	32 (1.1)	25 (0.9)	8 (0.6)	65 (1.0)	33 (1.0)
	2000	33 (1.2)	31 (1.5)	26 (1.0)	10 (0.8)	67 (1.2)	36 (1.2)

Standard errors of the estimated percentages appear in parentheses.
* Significantly different from 2000.
NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.9: Data for Figure 2.3 Scale Score Results by Race/Ethnicity

Percentage of students and average fourth-grade reading scale scores by race/ethnicity: 1992–2000

	White	Black	Hispanic	Asian/ Pacific Islander	American Indian
1992	71 (0.2) 225 (1.2)	16 (0.1) 193 (1.6)	9 (0.1) 201 (2.1)	2 (0.3) 214 (3.3) *	2 (0.2) 207 (4.6)
1994	69 (0.2) 224 (1.3)	15 (0.2) 187 (1.7) *	12 (0.2) 191 (2.6)	3 (0.2) 229 (4.4)	2 (0.2) 201 (3.4)
1998	67 (0.5) 227 (0.8)	16 (0.2) 194 (1.7)	13 (0.5) 196 (1.8)	2 (0.2) 225 (2.7)	2 (0.2) 202 (3.1)
2000	66 (0.4) 226 (1.0)	14 (0.2) 193 (1.7)	15 (0.3) 197 (1.7)	3 (0.2) 232 (4.6)	2 (0.2) 196 (4.7)

The percentage of students is listed first with the corresponding average scale score presented below.

Standard errors of the estimated percentages and scale scores appear in parentheses.

* Significantly different from 2000.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.10: Data for Figure 2.4 Achievement Level Results by Race/Ethnicity

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by race/ethnicity: 1992–2000

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
White	1992	29 (1.3)	36 (1.4)	27 (1.3)	8 (0.9)	71 (1.3)	35 (1.7)
	1994	29 (1.2)	34 (0.8)	27 (1.1)	9 (0.9)	71 (1.2)	37 (1.4)
	1998	27 (1.1)	34 (0.8)	29 (0.9)	10 (0.7)	73 (1.1)	39 (1.1)
	2000	27 (1.1)	33 (1.3)	29 (1.0)	11 (0.7)	73 (1.1)	40 (1.2)
Black	1992	67 (2.3)	25 (2.8)	8 (1.3)	1 (0.4)	33 (2.3)	8 (1.4)
	1994	69 (2.5)	22 (2.2)	8 (0.9)	1 (0.4)	31 (2.5)	9 (1.0)
	1998	64 (1.7)	26 (1.6)	9 (1.0)	1 (0.5)	36 (1.7)	10 (1.0)
	2000	63 (1.6)	25 (1.5)	10 (1.2)	2 (0.6)	37 (1.6)	12 (1.3)
Hispanic	1992	56 (2.2)	28 (2.0)	13 (1.6)	3 (0.8)	44 (2.2)	16 (1.8)
	1994	64 (2.6)	22 (2.1)	11 (1.5)	2 (0.6)	36 (2.6)	13 (1.6)
	1998	60 (1.9)	26 (1.5)	11 (1.2)	2 (0.4)	40 (1.9)	13 (1.2)
	2000	58 (1.8)	26 (1.6)	13 (1.4)	3 (0.6)	42 (1.8)	16 (1.3)
Asian/Pacific Islander	1992	41 (4.8)	34 (4.2)	21 (4.7)	4 (1.8)	59 (4.8)	25 (4.7) *
	1994	25 (3.9)	31 (4.0)	31 (3.5)	13 (4.9)	75 (3.9)	44 (5.5)
	1998	31 (4.2)	32 (4.7)	25 (3.4)	12 (2.9)	69 (4.2)	37 (3.9)
	2000	22 (5.0)	32 (3.9)	29 (3.8)	17 (4.6)	78 (5.0)	46 (4.4)
American Indian	1992	47 (6.6)	35 (6.5)	15 (4.5)	3 (2.1)	53 (6.6)	18 (4.5)
	1994	52 (4.4)	30 (4.2)	15 (3.8)	3 (2.1)	48 (4.4)	18 (3.8)
	1998	53 (5.5)	33 (5.2)	12 (4.5)	2 (1.2)	47 (5.5)	14 (3.8)
	2000	57 (4.2)	26 (6.1)	16 (4.6)	2 (****)	43 (4.2)	17 (4.8)

Standard errors of the estimated percentages appear in parentheses.

* Significantly different from 2000.

(****) Standard error estimates cannot be accurately determined.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.11: Data for Figure 2.5 Scale Score Differences Between Selected Subgroups

Differences in average fourth-grade reading scale scores by gender and race/ethnicity: 1992–2000

	Female/male	White/black	White/Hispanic
1992	8 (1.6)	32 (2.0)	23 (2.4)
1994	10 (1.7)	37 (2.1)	33 (2.9)
1998	6 (1.3) *	33 (1.9)	31 (2.0)
2000	10 (1.4)	33 (2.0)	29 (2.0)

Standard errors of the estimated difference in scale scores appear in parentheses.

* Significantly different from 2000.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.12: Data for Figure 2.6 Scale Score Results by Region

Percentage of students and average fourth-grade reading scale scores by regions of the country: 1992–2000

	Northeast	Southeast	Central	West
1992	21 (1.1) 222 (3.6)	23 (1.0) 213 (2.3)	27 (0.5) 219 (1.4)	28 (0.8) 214 (1.4)
1994	23 (0.9) 215 (2.1) *	23 (1.1) 210 (2.0)	25 (0.7) 220 (2.4)	29 (0.8) 212 (2.0)
1998	22 (1.0) 226 (1.4)	26 (1.1) 211 (1.3)	24 (0.6) 222 (2.0)	28 (1.3) 212 (1.9)
2000	22 (0.7) 222 (1.7)	22 (1.5) 211 (1.9)	25 (0.5) 220 (1.8)	31 (1.5) 214 (1.6)

The percentage of students is listed first with the corresponding average scale score presented below.

Standard errors of the estimated percentages and scale scores appear in parentheses.

* Significantly different from 2000.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.13: Data for Figure 2.7 Achievement Level Results by Region

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by regions of the country: 1992–2000

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Northeast	1992	34 (3.6)	32 (2.2)	24 (2.6)	9 (2.4)	66 (3.6)	34 (4.3)
	1994	39 (2.1)	30 (1.7)	23 (1.7)	8 (1.4)	61 (2.1)	31 (2.4)
	1998	30 (1.7)	32 (1.3)	28 (1.3)	10 (1.2)	70 (1.7)	38 (1.7)
	2000	33 (2.0)	30 (2.0)	26 (1.6)	11 (1.1)	67 (2.0)	37 (1.8)
Southeast	1992	42 (3.1)	34 (2.3)	19 (2.0)	5 (1.0)	58 (3.1)	24 (2.6)
	1994	45 (2.3)	30 (1.7)	19 (1.7)	7 (0.9)	55 (2.3)	25 (2.1)
	1998	44 (1.6)	31 (1.2)	20 (1.0)	5 (0.7)	56 (1.6)	25 (1.4)
	2000	43 (1.8)	31 (1.7)	20 (1.3)	6 (0.7)	57 (1.8)	26 (1.8)
Central	1992	34 (1.7)	36 (1.7)	24 (1.6)	6 (1.1)	66 (1.7)	30 (2.1)
	1994	34 (2.6)	33 (1.5)	26 (2.0)	8 (1.1)	66 (2.6)	34 (2.5)
	1998	32 (2.2)	34 (1.1)	26 (1.5)	8 (0.9)	68 (2.2)	35 (1.9)
	2000	34 (2.2)	31 (1.8)	26 (1.9)	8 (1.2)	66 (2.2)	35 (2.4)
West	1992	41 (1.7)	32 (1.5)	21 (1.7)	6 (0.7)	59 (1.7)	27 (1.7)
	1994	41 (2.1)	30 (1.5)	22 (1.5)	7 (0.8)	59 (2.1)	29 (1.8)
	1998	43 (2.3)	30 (1.9)	21 (1.5)	6 (0.8)	57 (2.3)	27 (2.0)
	2000	39 (1.3)	30 (1.3)	23 (1.3)	8 (0.9)	61 (1.3)	30 (1.3)

Standard errors of the estimated percentages appear in parentheses.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.14: Data for Table 2.1 Scale Score Results by Type of Location

Percentage of students and average reading scale scores by school's type of location: 2000

	Central city	Urban fringe/large town	Rural/small town
2000	32 (1.7) 209 (1.6)	45 (2.4) 222 (1.8)	23 (2.1) 218 (1.8)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Table C.15: Data for Figure 2.8 Achievement Level Results by Type of Location

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by school's type of location: 2000

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Central city	47 (1.9)	27 (1.6)	20 (1.1)	6 (0.9)	53 (1.9)	26 (1.7)
Urban fringe/large town	32 (1.6)	32 (1.7)	26 (1.2)	10 (1.1)	68 (1.6)	36 (1.8)
Rural/small town	35 (1.7)	33 (1.9)	25 (1.9)	8 (0.8)	65 (1.7)	32 (2.3)

Standard errors of the estimated percentages appear in parentheses.
NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Reading Assessment.

Table C.16: Data for Figure 2.9 Scale Score Results by Free/Reduced-Price Lunch Eligibility

Percentage of students and average fourth-grade reading scale scores by student eligibility for the free/reduced-price lunch program: 1998–2000

	Eligible	Not eligible	Information not available
1998	35 (1.2) 198 (1.2)	54 (1.8) 227 (0.9)	12 (1.9) 227 (2.8)
2000	34 (1.2) 196 (1.2)	51 (1.9) 227 (1.2)	15 (2.2) 228 (1.9)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table C.17: Data for Figure 2.10 Achievement Level Results by Free/Reduced-Price Lunch Eligibility

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by student eligibility for the free/reduced-price lunch program: 1998–2000

		Below Basic	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Eligible	1998	58 (1.4)	29 (1.4)	11 (1.1)	2 (0.4)	42 (1.4)	13 (1.2)
	2000	60 (1.3)	26 (1.2)	12 (0.9)	2 (0.3)	40 (1.3)	14 (0.9)
Not Eligible	1998	27 (1.2)	33 (1.2)	30 (1.0)	10 (0.9)	73 (1.2)	40 (1.3)
	2000	26 (1.2)	34 (1.4)	30 (1.0)	11 (0.8)	74 (1.2)	41 (1.3)
Information not available	1998	27 (2.7)	33 (2.4)	29 (3.1)	11 (1.6)	73 (2.7)	40 (3.8)
	2000	26 (2.0)	32 (1.9)	30 (2.1)	12 (1.3)	74 (2.0)	42 (2.1)

Standard errors of the estimated percentages appear in parentheses.
NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table C.18: Data for Figure 2.11 Scale Score Results by Type of School

Percentage of students and average fourth-grade reading scale scores by type of school: 1992–2000

	Public	Nonpublic	Nonpublic: Catholic	Other Nonpublic
1992	89 (1.0)	11 (1.0)	8 (0.8)	4 (0.7)
	215 (1.0)	232 (1.7)	229 (2.2)	238 (2.9) !
1994	90 (0.9)	10 (0.9)	7 (0.8)	4 (0.6)
	212 (1.1)	231 (2.5)	229 (3.3)	234 (3.7)
1998	89 (1.2)	11 (1.2)	7 (1.0)	4 (0.6)
	215 (0.8)	233 (2.3)	233 (2.5)	232 (4.5)
2000	89 (0.7)	11 (0.7)	6 (0.5)	5 (0.5)
	215 (0.9)	234 (1.7)	231 (2.6)	237 (2.1)

The percentage of students is listed first with the corresponding average scale score presented below.

Standard errors of the estimated percentages and scale scores appear in parentheses.

! The nature of the sample does not allow accurate determination of the variability of the statistic.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

BEST COPY AVAILABLE

Table C.19: Data for Figure 2.12 Achievement Level Results by Type of School

Percentages of fourth-grade students at or above reading achievement levels and within each achievement level range by type of school: 1992–2000

		Below Basic	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Public	1992	40 (1.1)	33 (1.0)	21 (1.0)	6 (0.6)	60 (1.1)	27 (1.3)
	1994	41 (1.1)	30 (0.8)	21 (1.0)	7 (0.7)	59 (1.1)	28 (1.2)
	1998	39 (1.0)	31 (0.8)	23 (0.8)	6 (0.5)	61 (1.0)	29 (0.9)
	2000	40 (0.9)	31 (1.0)	22 (0.8)	7 (0.6)	60 (0.9)	30 (1.0)
Nonpublic	1992	21 (1.9)	34 (2.2)	33 (2.2)	12 (1.3)	79 (1.9)	45 (2.4)
	1994	23 (2.4)	34 (2.2)	31 (2.0)	13 (1.8)	77 (2.4)	43 (3.0)
	1998	22 (2.6)	32 (2.1)	32 (2.1)	14 (1.5)	78 (2.6)	46 (2.9)
	2000	20 (1.8)	32 (1.8)	34 (1.8)	14 (1.4)	80 (1.8)	47 (2.4)
Nonpublic: Catholic	1992	24 (2.7)	35 (2.2)	30 (2.0)	10 (1.5)	76 (2.7)	41 (2.7)
	1994	24 (3.2)	34 (2.9)	30 (2.8)	12 (2.2)	76 (3.2)	42 (3.9)
	1998	21 (2.9)	33 (2.5)	32 (2.5)	13 (1.7)	79 (2.9)	46 (3.3)
	2000	22 (2.9)	33 (2.2)	33 (2.3)	11 (1.9)	78 (2.9)	44 (3.1)
Other Nonpublic	1992	16 (2.7) !	31 (4.1) !	38 (4.5) !	15 (2.9) !	84 (2.7) !	53 (4.4) !
	1994	20 (4.2)	34 (3.0)	32 (2.6)	14 (2.9)	80 (4.2)	46 (4.0)
	1998	24 (4.6)	30 (3.0)	31 (3.6)	16 (2.9)	76 (4.6)	46 (5.0)
	2000	18 (2.1)	31 (2.7)	35 (2.3)	16 (1.8)	82 (2.1)	51 (2.9)

Standard errors of the estimated percentages appear in parentheses.

! The nature of the sample does not allow accurate determination of the variability of the statistic.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.20: Data for Table 3.1 Pages Read in School and For Homework

Students' reports on the number of pages read each day in school and for homework: 1992–2000

	1992	1994	1998	2000
11 or more pages	56 (1.2) *	54 (1.1) *	57 (1.1)	60 (1.2)
	222 (1.1)	220 (1.3)	221 (0.8)	222 (0.9)
6 to 10 pages	23 (0.7) *	23 (0.7) *	22 (0.6)	20 (0.8)
	217 (1.3)	214 (1.3)	217 (1.3)	215 (1.5)
5 or fewer pages	21 (1.0)	23 (0.8) *	21 (0.8)	19 (0.8)
	203 (1.4)	201 (1.2)	207 (1.3)	202 (2.1)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 2000.
NOTE: Percentages may not add up to 100 due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.21: Data for Table 3.2 Time Spent Doing Homework

Students' reports on the amount of time spent doing homework each day: 1992–2000

	1992	1994	1998	2000
More than one hour	15 (0.6)	15 (0.6)	16 (0.6)	16 (0.7)
	208 (1.9)	208 (2.1)	213 (1.7)	212 (1.7)
One hour	28 (0.9)	30 (0.7)	31 (0.8)	29 (0.6)
	221 (1.1)	218 (1.4)	221 (1.2)	222 (1.2)
One-half hour	39 (1.2) *	39 (1.0) *	43 (1.0)	43 (0.9)
	217 (1.3)	216 (1.3)	219 (1.1)	219 (1.1)
Do not do homework	2 (0.2)	3 (0.3) *	2 (0.2)	2 (0.2)
	196 (3.7)	183 (3.0)	188 (3.4)	172 (4.1)
Do not have homework	16 (1.6) *	13 (0.9) *	8 (0.8)	10 (0.9)
	220 (1.6)	216 (2.2)	213 (2.7)	212 (3.0)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 2000.
NOTE: Percentages may not add up to 100 due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

BEST COPY AVAILABLE

Table C.22: Data for Table 3.3 Writing About Reading

Students' reports on how often they write long answers to questions on tests or assignments that involved reading: 1992–2000

	1992	1994	1998	2000
At least once a week	51 (1.0) 220 (1.1)	48 (0.8) * 217 (1.0)	53 (0.8) 218 (0.8)	53 (1.1) 217 (1.0)
Once or twice a month	28 (0.8) 221 (1.2)	31 (0.7) * 221 (1.4)	30 (0.6) 223 (1.0)	28 (0.9) 225 (1.2)
Once or twice a year	13 (0.5) 209 (2.2)	12 (0.5) 209 (2.2)	10 (0.5) 212 (1.8)	11 (0.4) 210 (2.5)
Never or hardly ever	9 (0.5) 202 (2.1)	9 (0.4) * 198 (2.8)	8 (0.4) 199 (2.1)	8 (0.4) 199 (2.6)

The percentage of students is listed first with the corresponding average scale score presented below.

Standard errors of the estimated percentages and scale scores appear in parentheses.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.23: Data for Table 3.4 Teachers' Help with Words

Students' reports on how often their teachers help them break words into parts and help them understand new words: 1998–2000

	1998	2000
How often their teachers help them break words into parts		
Every day	25 (0.7) 210 (1.2)	25 (0.9) 209 (1.3)
Once or twice a week	23 (0.6) 217 (1.3)	22 (0.7) 217 (1.2)
Never or hardly ever	52 (0.8) 226 (1.0)	53 (0.9) 226 (1.1)
How often their teachers help them understand new words		
Every day	49 (0.9) 217 (0.9)	48 (0.9) 217 (1.1)
Once or twice a week	24 (0.7) 224 (1.2)	23 (0.6) 224 (1.1)
Once or twice a month	14 (0.5) 223 (1.9)	14 (0.5) 224 (1.9)
Never or hardly ever	12 (0.5) * 219 (1.8)	14 (0.5) 216 (2.1)

The percentage of students is listed first with the corresponding average scale score presented below.

Standard errors of the estimated percentages and scale scores appear in parentheses.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table C.24: Data for Table 3.5 Reading for Fun

Students' reports on how often they read for fun on their own time: 1992–2000

	1992	1994	1998	2000
Every day	44 (0.9) 223 (1.2)	45 (0.7) 223 (1.2)	43 (0.7) 222 (1.1)	43 (0.9) 223 (1.1)
Once or twice a week	32 (0.8) 218 (1.2)	32 (0.7) 213 (1.1)	32 (0.6) 219 (1.0)	32 (0.7) 218 (0.9)
Once or twice a month	12 (0.4) 210 (1.6)	12 (0.5) 208 (2.1)	12 (0.6) 216 (1.7)	12 (0.5) 216 (1.6)
Never or hardly ever	13 (0.5) 199 (1.8)	12 (0.4) * 197 (1.9)	13 (0.5) 203 (1.4)	14 (0.5) 202 (2.8)

The percentage of students is listed first with the corresponding average scale score presented below.

Standard errors of the estimated percentages and scale scores appear in parentheses.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.25: Data for Table 3.6 Discussing Studies and Talking About Reading

Students' reports on how often they discuss their studies at home and talk about reading with their family and friends: 1992–2000

	1992	1994	1998	2000
Discuss studies at home				
Almost every day	54 (0.8)	55 (0.8)	54 (0.6)	54 (0.7)
	221 (1.0)	219 (1.0)	220 (0.8)	221 (1.1)
Once or twice a week	22 (0.7)	22 (0.5)	23 (0.6)	23 (0.6)
	220 (1.5)	215 (1.7)	222 (1.3)	219 (1.2)
Once or twice a month	6 (0.3)	6 (0.4)	6 (0.3)	6 (0.3)
	215 (1.8)	208 (2.3)	213 (2.2)	217 (3.5)
Never or hardly ever	17 (0.6)	17 (0.5)	18 (0.5)	17 (0.5)
	202 (1.5)	199 (1.7)	205 (1.3)	201 (1.7)
Talk about reading with family or friends				
Almost every day	26 (0.6)	28 (0.6)	27 (0.5)	27 (0.6)
	215 (1.4)	213 (1.3)	211 (1.2)	213 (1.6)
Once or twice a week	36 (0.9)	36 (0.6)	35 (0.7)	34 (0.7)
	224 (1.1)	223 (1.2)	223 (1.0)	227 (1.0)
Once or twice a month	15 (0.6)	15 (0.5)	15 (0.5)	15 (0.5)
	219 (1.6)	214 (2.1)	222 (1.5)	220 (1.8)
Never or hardly ever	23 (0.8)	21 (0.6) *	23 (0.7)	24 (0.6)
	209 (1.4)	207 (1.6)	214 (1.3)	209 (1.1)

The percentage of students is listed first with the corresponding average scale score presented below.

Standard errors of the estimated percentages and scale scores appear in parentheses.

* Significantly different from 2000.

NOTE: Percentages may not add up to 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.26: Data for Table 3.7 Reading Materials in the Home

Students' reports on the number of different types of reading materials in the home: 1992–2000

	1992	1994	1998	2000
Four	37 (0.9) 226 (1.3)	38 (0.8) * 227 (1.1)	37 (1.0) 228 (1.1)	34 (0.8) 229 (1.3)
Three	32 (0.7) 219 (1.3)	34 (0.7) 216 (1.2)	33 (0.7) 220 (1.1)	34 (0.8) 219 (1.0)
Two or fewer	31 (0.8) 204 (0.9)	29 (0.9) * 197 (1.4)	30 (0.8) 204 (1.1)	32 (1.0) 203 (1.3)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 2000.
NOTE: Percentages may not add up to 100 due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.27: Data for Table 3.8 Time Spent Watching Television

Students' reports on the amount of time spent watching television each day: 1992–2000

	1992	1994	1998	2000
Six hours or more	20 (0.7) 199 (1.5)	21 (0.7) * 194 (1.4)	16 (0.6) * 198 (1.5)	18 (0.6) 196 (1.7)
Four or five hours	22 (0.8) * 216 (1.3)	22 (0.7) * 216 (1.7)	19 (0.7) * 216 (1.4)	17 (0.6) 213 (2.2)
Two or three hours	40 (0.8) 224 (1.0)	38 (0.7) * 222 (1.1)	41 (0.9) 223 (0.9)	40 (0.7) 224 (1.0)
One hour or less	19 (0.8) * 221 (1.6)	19 (0.7) * 220 (1.9)	24 (0.7) 222 (1.3)	25 (0.8) 224 (1.4)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 2000.
NOTE: Percentages may not add up to 100 due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, and 2000 Reading Assessments.

Table C.28: Data for Table 4.1

Average score by type of results: 1998–2000

	Accommodations not permitted	*Accommodations permitted*
1998	217 (0.8)	216 (0.9)
2000	217 (0.8)	215 (1.0) †

Standard errors of the estimated scale scores appear in parentheses.
†Significantly different from the sample where accommodations were not permitted.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table C.29: Data for Table 4.2

Percentages of students attaining reading achievement levels by type of results: 1998–2000

	Below Basic	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
1998						
Accommodations were						
not permitted	38 (0.9)	32 (0.7)	24 (0.7)	7 (0.5)	62 (0.9)	31 (0.9)
permitted	39 (1.0)	31 (0.8)	23 (0.8)	8 (0.5)	61 (1.0)	31 (0.9)
2000						
Accommodations were						
not permitted	37 (0.8)	31 (0.9)	24 (0.8)	8 (0.5)	63 (0.8)	32 (0.9)
permitted	39 (1.1)	30 (0.9)	23 (0.9)	7 (0.6)	61 (1.1)	31 (0.9)

Standard errors of the estimated percentages appear in parentheses.
NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table C.30

Average scores by gender and type of results (Accommodations Not Permitted and Accommodations Permitted): 1998–2000

	Male		Female	
	Not permitted	*Permitted*	*Not Permitted*	*Permitted*
1998	214 (1.1)	214 (1.2)*	220 (0.7)	219 (1.1)
2000	212 (1.1)	210 (1.0)	222 (0.9)	220 (1.2)

Standard errors of the estimated scale scores appear in parentheses.
*Significantly different from 2000.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table C.31

Percentages of students attaining reading achievement levels by gender and by type of results: 1998–2000

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Male						
1998: Accommodations were						
not permitted	41 (1.4)	31 (1.0)	22 (0.9)	6 (0.6)	59 (1.4)	28 (1.2)
permitted	41 (1.2)	30 (1.0)	22 (1.0)	6 (0.7)	59 (1.2)	28 (1.1)
2000: Accommodations were						
not permitted	42 (1.2)	31 (1.1)	21 (1.2)	6 (0.5)	58 (1.2)	27 (1.1)
permitted	44 (1.2)	30 (1.4)	20 (1.0)	6 (0.6)	56 (1.2)	26 (1.1)
Female						
1998: Accommodations were						
not permitted	35 (1.0)	32 (1.1)	25 (0.9)	8 (0.6)	65 (1.0)	33 (1.0)
permitted	36 (1.2)	31 (1.0)	24 (1.1)	9 (0.8)	64 (1.2)	33 (1.1)
2000: Accommodations were						
not permitted	33 (1.2)	31 (1.5)	26 (1.0)	10 (0.8)	67 (1.2)	36 (1.2)
permitted	34 (1.3)	30 (1.0)	26 (1.2)	9 (1.0)	66 (1.3)	35 (1.2)

Standard errors of the estimated percentages appear in parentheses.
NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Table C.32

Average scores by race/ethnicity and type of results (Accommodations Not Permitted and Accommodations Permitted): 1998–2000

	White		Black		Hispanic		Asian Pacific Islander		American Indian	
	Not permitted	*Permitted*	*Not permitted*	*Permitted*	*Not permitted*	*Permitted*	*Not permitted*	*Permitted*	*Not permitted*	*Permitted*
1998	227 (0.8)	226 (1.0)	194 (1.7)	194 (1.[illegible])	196 (1.8)	193 (2.5)	225 (2.7)	220 (3.8)	[illegible] (3.1)	199 (3.0)
2000	226 (1.0)	225 (1.0)	193 (1.7)	193 (1.[illegible])	197 (1.7)	190 (2.5) †	232 (4.6)	229 (4.3)	196 (4.7)	201 (4.3)

Standard errors of the estimated scale scores appear in parentheses.
†Significantly different from the sample where accommodations were not permitted.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

BEST COPY AVAILABLE

Table C.33

Percentages of students attaining reading achievement levels by race/ethnicity and by type of results: 1998–2000

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
White						
1998: Accommodations were						
not permitted	27 (1.1)	34 (0.8)	29 (0.9)	10 (0.7)	73 (1.1)	39 (1.1)
permitted	28 (1.3)	33 (1.1)	29 (1.2)	10 (0.8)	72 (1.3)	39 (1.2)
2000: Accommodations were						
not permitted	27 (1.1)	33 (1.3)	29 (1.0)	11 (0.7)	73 (1.1)	40 (1.2)
permitted	29 (1.3)	33 (1.2)	29 (1.2)	10 (0.8)	71 (1.3)	39 (1.2)
Black						
1998: Accommodations were						
not permitted	64 (1.7)	26 (1.6)	9 (1.0)	1 (0.5)	36 (1.7)	10 (1.0)
permitted	63 (2.1)	26 (2.0)	9 (1.2)	1 (0.4)	37 (2.1)	10 (1.1)
2000: Accommodations were						
not permitted	63 (1.6)	25 (1.5)	10 (1.2)	2 (0.6)	37 (1.6)	12 (1.3)
permitted	63 (1.7)	26 (1.4)	9 (1.1)	1 (0.6)	37 (1.7)	11 (1.1)
Hispanic						
1998: Accommodations were						
not permitted	60 (1.9)	26 (1.5)	11 (1.2)	2 (0.4)	40 (1.9)	13 (1.2)
permitted	63 (2.5)	25 (1.8)	11 (1.2)	2 (0.4)	37 (2.5)	13 (1.3)
2000: Accommodations were						
not permitted	58 (1.8)	26 (1.6)	13 (1.4)	3 (0.6)	42 (1.8)	16 (1.3)
permitted	62 (2.3)	24 (2.4)	12 (1.3)	2 (0.5)	38 (2.3)	14 (1.4)
Asian/Pacific Islander						
1998: Accommodations were						
not permitted	31 (4.2)	32 (4.7)	25 (3.4)	12 (2.9)	69 (4.2)	37 (3.9)
permitted	37 (5.0)	29 (4.4)	23 (3.4)	11 (2.8)	63 (5.0)	34 (3.7)
2000: Accommodations were						
not permitted	22 (5.0)	32 (3.9)	29 (3.8)	17 (4.6)	78 (5.0)	46 (4.4)
permitted	26 (4.4)	30 (4.8)	28 (4.0)	16 (3.0)	74 (4.4)	44 (5.0)
American Indian						
1998: Accommodations were						
not permitted	53 (5.5)	33 (5.2)	12 (4.5)	2 (1.2)	47 (5.5)	14 (3.8)
permitted	58 (5.4)	29 (4.3)	12 (3.5)	1 (0.6)	42 (5.4)	13 (3.2)
2000: Accommodations were						
not permitted	57 (4.2)	26 (6.1)	16 (4.6)	2 (****)	43 (4.2)	17 (4.8)
permitted	51 (5.1)	29 (6.9)	18 (4.6)	2 (****)	49 (5.1)	20 (4.8)

Standard errors of the estimated percentages appear in parentheses.

(****) Standard error estimates cannot be accurately determined.

NOTE: Percentages within each reading achievement level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2000 Reading Assessments.

Acknowledgments

This report is the culmination of the effort of many individuals who contributed their considerable knowledge, experience, and creativity to the NAEP 2000 reading assessment. The assessment was a collaborative effort among staff from the National Center for Education Statistics (NCES), the National Assessment Governing Board (NAGB), Educational Testing Service (ETS), Westat, and National Computer Systems (NCS). Most importantly, NAEP is grateful to the students and school staff who made the assessment possible.

The 2000 NAEP reading assessment was funded through NCES, in the Office of Educational Research and Improvement of the U.S. Department of Education. The acting Commissioner of Education Statistics, Gary W. Phillips, and the NCES staff—Peggy Carr, Patricia Dabbs, Arnold Goldstein, Carol Johnson, Andrew Kolstad, Holly Spurlock, and Sheida White—worked closely and collegially with the authors to produce this report.

The NAEP project at ETS is directed by Stephen Lazer and John Mazzeo. Sampling and data collection activities were conducted by Westat under the direction of Rene Slobasky, Nancy Caldwell, Keith Rust, and Dianne Walsh. Printing, distribution, scoring, and processing activities were conducted by NCS under the direction of Brad Thayer, Patrick Bourgeacq, William Buckles, Mathilde Kennel, Linda Reynolds, and Connie Smith.

The complex statistical and psychometric activities necessary to report results for the NAEP 2000 reading assessment were directed by Nancy Allen with assistance from Spence Swinton. The numerous computer programs associated with these activities were carried out by Steve Isham and Lois Worthington. The analyses presented in this report were produced by Dave Freund, Steve Isham, Laura Jerry, Youn-hee Lim, and Lois Worthington.

The design and production of this report was overseen by Loretta Casalaina. Joseph Kolodey and Carol Erickson contributed invaluable design and production expertise to the effort. Assistance in the production of the reports was also provided by Barbette Tardugno. Wendy Grigg coordinated the documentation and data checking procedures with assistance from Alice Kass. Shari Santapau coordinated the editorial and proofreading procedures. The web version of this report was coordinated by Pat O'Reilly with assistance from Rick Hasney. These activities, as well as statistical and psychometric activities, were conducted at ETS.

Many thanks are due to the numerous reviewers, both internal and external to NCES and ETS. The comments and critical feedback of the following reviewers are reflected in the final version of this report: Peter Afflerbach, Mary Crovo, Patricia Dabbs, Lawrence Feinberg, Steve Gorman, David Grissmer, Carol Johnson, Janet Johnson, Andrew Kolstad, Marilyn McMillen, Holly Spurlock, Alan Vanneman, and Sheida White.

United States
Department of Education
ED Pubs
8242-B Sandy Court
Jessup, MD 20794-1398

Official Business Only
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G-17

Standard Mail (B)